



University
of Glasgow

Jones, Nathaniel Gadsby (2013) *Validating protein kinases of Trypanosoma brucei as drug targets*. PhD thesis.

<http://theses.gla.ac.uk/4007/>

Copyright and moral rights for this thesis are retained by the author

A copy can be downloaded for personal non-commercial research or study

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the Author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the Author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Validating Protein Kinases of *Trypanosoma brucei* as Drug Targets

Nathaniel Gadsby Jones

BSc (Hons), MSc (Dist)

**Thesis submitted in fulfilment of the requirements
for the degree of Doctor of Philosophy**

**Institute of Infection, Immunity and Inflammation
College of Medical, Veterinary and Life Sciences**

University of Glasgow

February 2013

Abstract

Trypanosoma brucei spp. are protozoan parasites that cause Human African Sleeping Sickness and Nagana, a disease of cattle. The diseases have a very high mortality rate if untreated and the current drug treatments are inadequate due to toxicity and resistance. In order to develop new treatments, potential drug targets can be investigated in a candidate approach by genetic validation. In trypanosomes this can be performed by using RNA interference (RNAi) technology, which is well developed for this organism. One category of potential targets are protein kinases; members of this enzyme family have been shown to be suitable targets for relatively specific and selective, therapeutic inhibition- especially in the field of human oncology. *T. brucei* possesses 183 eukaryotic protein kinases, atypical protein kinase and pseudokinases that may be suitable for chemical inhibition. This study intended to genetically validated these targets and pursue interesting leads as potential drug targets.

In order to perform RNAi studies on large numbers of candidate genes an existing stem-loop RNAi system was adapted to contain Gateway recombination sites. This created a highly efficient system, allowing the rapid generation of RNAi plasmids to target the entire complement of the predicted kinome of this organism. This collection of plasmids was used to generate a library of RNAi cell lines in bloodstream form parasites (BSF) which were screened for growth defects using an alamar blue cell viability assay. This identified 53 genes which were essential or important for proliferation of the BSF parasites.

During this screen, a cell line was identified that contained an RNAi construct targeting two NEK kinases (NEK12.1: Tb927.8.7110, NEK12.2: Tb927.4.5310) which displayed a severe growth defect. This was replicated in a mouse model of infection. NEK 12.1 possesses an alanine gatekeeper residue and molecular modelling and virtual screening predicted this would allow the accommodation of bulky protein kinase inhibitors and therefore was investigated in more detail. Expression and purification of NEK12.1 in kinase active and dead forms allowed activity assays to be performed; these could be inhibited with the bulky inhibitor 3-MB-PP1, which also possessed activity against BSF parasites. Experiments to confirm the individual essentiality of NEK 12.1 (and its activity) remain to be performed, therefore NEK 12.1 has been partially genetically and chemically

validated in this study. A unique “Orphan” kinase, containing a putative zinc-finger domain and a potential homologue of PDK1 were also investigated after the pTL RNAi screen. *In vivo* RNAi studies (using a mouse model) correlated well with the in culture RNAi data, demonstrating that these targets are essential in mammalian infections.

During the pTL RNAi screen it was noted that the knockdown of one protein kinase, (DFK: Tb11.01.5650) led to a change in the morphology of the cells, yet no reduction in the alamar blue ratio was observed. Further investigation showed that after 72 h of RNAi induction 17% of cells expressed EP procyclin. This was associated with detectable changes in the transcriptome (ascertained by RNA -Seq), that were consistent with a BSF-PCF differentiation event. DFK was predicted to contain multiple transmembrane domains and features suggestive of a receptor kinase. Epitope tagging of DFK followed by cell fractionation and immunofluorescence microscopy demonstrated that the protein localised to the cell membrane. Transmission immuno-electron microscopy confirmed the cell membrane localisation and suggested that the protein kinase domain of DFK was intracellular.

This study reveals several protein kinases as new drug targets, in turn identifying one as a key regulator of BSF-PCF differentiation. Due to the tractability of *T. brucei* it also provides a collection of plasmids and cell lines that should further the investigation of other key biological processes in this important pathogen.

Author's Declaration

The results stated in this thesis are my own work, except where otherwise stated, and have not been submitted for any other degree.

Nathaniel Gadsby Jones

Table of Contents

ABSTRACT	1
AUTHOR'S DECLARATION	3
LIST OF TABLES.....	7
LIST OF FIGURES	8
LIST OF ACCOMPANYING MATERIAL	10
ABBREVIATIONS.....	11
AMINO ACID ABBREVIATIONS.....	13
ACKNOWLEDGEMENTS	14
1 GENERAL INTRODUCTION	16
1.1 AFRICAN TRYPANOSOMIASES CAUSED BY <i>TRYPANOSOMA BRUCEI</i> SPP.	16
1.2 <i>TRYPANOSOMA BRUCEI</i> LIFE HISTORY, CELLULAR AND MOLECULAR BIOLOGY	22
1.3 PROTEIN KINASES	28
1.4 PROTEIN KINASES AS DRUG TARGETS.....	31
1.5 PROTEIN KINASE OF <i>T. BRUCEI</i>	33
1.6 VALIDATING DRUG TARGETS	39
1.7 RESEARCH AIMS	44
2 MATERIALS AND METHODS	45
2.1 BIOINFORMATICS	45
2.1.1 <i>Sequence Retrieval</i>	45
2.1.2 <i>Sequence Manipulation</i>	45
2.1.3 <i>Molecular Modelling</i>	45
2.1.4 <i>Data Mining for Potential NEK 12 Inhibitors</i>	46
2.1.5 <i>Virtual Screening</i>	46
2.2 MOLECULAR BIOLOGY	47
2.2.1 <i>DNA Sequencing</i>	47
2.2.2 <i>Polymerase Chain Reaction</i>	47
2.2.3 <i>Restriction Enzyme Digests</i>	48
2.2.4 <i>Agarose Gel Electrophoresis</i>	48
2.2.5 <i>Ligations</i>	49
2.2.6 <i>Gateway Cloning of RNAi Constructs</i>	49
2.2.7 <i>Ligation Independent Cloning of Kinase CDS for Protein Expression</i>	49
2.2.8 <i>Site Directed Mutagenesis</i>	59
2.2.9 <i>Transformation of Bacteria</i>	59
2.2.10 <i>Genomic DNA Extractions</i>	59
2.2.11 <i>Preparation of DNA for Trypanosome Transfection</i>	59
2.2.12 <i>RNA Extraction, cDNA Preparation, and qPCR</i>	60
2.3 PROTEIN PRODUCTION AND BIOCHEMISTRY	61
2.3.1 <i>Bacterial Cell Lines</i>	61
2.3.2 <i>Protein Expression</i>	62
2.3.3 <i>SDS-PAGE</i>	63
2.3.4 <i>Coomassie Staining of SDS-PAGE Gels</i>	63
2.3.5 <i>Western Blotting</i>	63
2.3.6 <i>Protein Kinase Assays</i>	64
2.4 TRYPANOSOME CELL CULTURE	64
2.4.1 <i>Cell Lines and General Culturing Techniques</i>	64
2.4.2 <i>Genetic Modification of Trypanosomes</i>	65
2.4.3 <i>Growth Curves</i>	66
2.4.4 <i>Mouse Infections</i>	66
2.4.5 <i>Alamar Blue Screening for RNAi Phenotypes</i>	66
2.5 FURTHER ANALYSES.....	67
2.5.1 <i>Flow Cytometry for DNA Content and EP Procyclin</i>	67

2.5.2	<i>Immunofluorescence for Epitope Tagged DFK</i>	67
2.5.3	<i>Cell Fractionation</i>	68
2.5.4	<i>Transmission Immuno-Electron Microscopy</i>	69
2.5.5	<i>Immunoprecipitation of DFK</i>	69
3	GENERATION AND ANALYSIS OF A KINOME-WIDE RNAI LIBRARY IN <i>TRYPANOSOMA BRUCEI</i>	71
3.1	INTRODUCTION.....	71
3.1.1	<i>RNA Interference in T. brucei</i>	71
3.1.2	<i>RNAi as a Tool for Investigating Trypanosome Biology</i>	72
3.1.3	<i>RNAi Screens in T. brucei</i>	73
3.2	RESEARCH AIMS	77
3.3	RESULTS.....	77
3.3.1	<i>Generation and Validation of the pTL System</i>	77
3.3.2	<i>Validation of the System</i>	80
3.3.3	<i>Generation of kinome-wide RNAi plasmid library</i>	81
3.3.4	<i>Generation of RNAi-cell lines</i>	82
3.3.5	<i>Alamar blue screening for growth defects under RNAi induction</i>	82
3.3.6	<i>Results of Kinome-Wide RNAi Screen</i>	84
3.3.7	<i>Validation of mRNA Knockdown by qPCR</i>	100
3.4	DISCUSSION	101
3.4.1	<i>Comparison of pTL Screen to Candidate Studies</i>	101
3.4.2	<i>Comparison to RIT-Seq Data</i>	103
3.4.3	<i>False Negatives</i>	104
3.4.4	<i>Other high-throughput cloning systems</i>	105
3.4.5	<i>Further Utilities of the RNAi Library</i>	105
4	<i>TRYPANOSOMA BRUCEI</i> PROTEIN KINASES AS DRUG TARGETS	110
4.1	TARGETS THAT WERE IDENTIFIED IN THE SCREEN FOR FURTHER STUDY	110
4.2	INTRODUCTION.....	110
4.2.1	<i>NEK Family Protein Kinases</i>	110
4.2.2	<i>PDK1</i>	115
4.2.3	<i>Orphan Kinases</i>	117
4.2.4	<i>Research Aims</i>	117
4.3	RESULTS.....	117
4.3.1	<i>Features of NEK12</i>	117
4.3.2	<i>NEK 12 RNAi</i>	119
4.3.3	<i>NEK 12 Phenotype – cell cycle analysis</i>	123
4.3.4	<i>Generation of Recombinant NEK 12 Proteins</i>	126
4.3.5	<i>Activity assays/inhibitors of NEKs</i>	129
4.3.6	<i>Homology Modelling of NEKs and Virtual Screening</i>	132
4.4	ZINC-FINGER CONTAINING ORPHAN KINASE.....	135
4.4.1	<i>Features of ZFO</i>	135
4.4.2	<i>RNAi Analysis</i>	136
4.5	PDK1.....	139
4.5.1	<i>Features of TbPDK1</i>	139
4.5.2	<i>RNAi of TbPDK1</i>	139
4.6	DISCUSSION	141
4.6.1	<i>NEK kinases as drug targets</i>	141
4.6.2	<i>ZFO and PDK1 as Drug Targets</i>	151
5	DFK – A NEGATIVE REGULATOR OF BLOODSTREAM FORM TO PROCYCLIC FORM DIFFERENTIATION 153	
5.1	INTRODUCTION.....	153
5.1.1	<i>Trypanosome Differentiation</i>	153
5.1.2	<i>Reversible phosphorylation during trypanosome differentiation</i>	156
5.1.3	<i>Research Aims</i>	162
5.2	RESULTS.....	162
5.2.1	<i>Identification and Features of DFK</i>	162
5.2.2	<i>RNAi Results</i>	164
5.2.3	<i>Evidence of Differentiation</i>	166

5.2.4	<i>Localisation of DFK</i>	170
5.3	DISCUSSION	184
5.3.1	<i>Confirming the DFK Phenotype</i>	184
5.3.2	<i>Localisation and Topology of DFK</i>	189
5.3.3	<i>Functional Role of DFK</i>	190
5.3.4	<i>Summary and Future Direction</i>	194
6	CONCLUDING REMARKS	197
	LIST OF REFERENCES	201

List of Tables

TABLE 1-1: TARGET-PRODUCT PROFILE FOR NEXT GENERATION ANTITRYPANOSOMALS AS DETERMINED BY DNDI.....	43
TABLE 2-1: TEMPLATE PDB STRUCTURES USED FOR HOMOLGY MODELLING OF <i>T. BRUCEI</i> NEK 12.1 AND NEK12.2	46
TABLE 2-2: LIST OF OLIGONUCLEOTIDES USED TO GENERATE RNAI CONSTRUCTS PRIOR TO THE ESTABLISHMENT OF A GATEWAY ADAPTED CLONING SYSTEM.	51
TABLE 2-3: LIST OF RNAI PLASMIDS GENERATED PRIOR TO THE GENERATION OF A GATEWAY ADAPTED VECTOR SYSTEM.....	52
TABLE 2-4: LIST OF OLIGONUCLEOTIDE PRIMERS GENERATED IN THIS STUDY FOR THE PURPOSES OF LIGATION INDEPENDENT CLONING INTO THE PET-30/32 XA/LIC SYSTEMS.....	53
TABLE 2-5: LIST OF OLIGONUCLEOTIDES GENERATED FOR SITE DIRECTED MUTAGENESIS OF CATALYTIC LYSINE RESIDUES IN VARIOUS PROTEIN KINASES.	56
TABLE 2-6: LIST OF PLASMIDS GENERATED FOR THE PURPOSE OF PROTEIN EXPRESSION IN E. COLI DE3 LYSOGEN CELL LINES.	56
TABLE 2-7: LIST OF OLIGONUCLEOTIDES USED TO GENERATE FRAGMENTS FOR CLONING OF ENDOGENOUS TAGGING VECTORS... ..	58
TABLE 2-8: LIST OF PLASMIDS GENERATED FOR ENDOGENOUS TAGGING OF GENES IN THIS STUDY.....	58
TABLE 2-9: LIST OF OLIGONUCLEOTIDES USED TO PRIME QPCR REACTIONS.....	61
TABLE 3-1: PROTEIN KINASES DETERMINED TO HAVE A SEVERE RNAI LOSS OF FITNESS PHENOTYPE AS DETERMINED BY ALAMAR BLUE RATIO.....	85
TABLE 3-2: PROTEIN KINASES DETECTED TO HAVE AN RNAI LOSS OF FITNESS PHENOTYPE AS DETERMINED BY AN ALAMAR BLUE ASSAY RATION BETWEEN 0.5 AND 0.9.....	87
TABLE 3-3: PROTEIN KINASES FOR WHICH NO LOSS OF FITNESS PHENOTYPE WAS DETECTED BY ALAMAR BLUE ASSAY.	89
TABLE 3-4: LIST OF RNAI PLASMIDS TARGETING MULTIPLE GENES.	98
TABLE 3-5: TABLE OF PTL PHENOTYPES CONFIRMING PREVIOUSLY PUBLISHED WORK ON BSF TRYPANOSOMES.	102
TABLE 4-1: VARIABLES TESTED TO OPTIMISE PROTEIN EXPRESSION,	127
TABLE 4-2: DETAILS OF WELLS IN FIGURE 4-9.	128
TABLE 5-1: LIST OF THE TRANSMEMBRANE PREDICTION ALGORITHMS USED TO ASSESS THE POTENTIAL TM DOMAINS IN DFK... ..	164
TABLE 5-2: OVERVIEW OF GENES DETECTED AS DIFFERENTIALLY REGULATED IN DFK RNAI CELLS AND THE EXISTING DATA AVAILABLE IN TriTRYPDB.....	181
TABLE 5-3: SELECTED TRANSCRIPTS IDENTIFIED AS DIFFERENTIALLY REGULATED IN DFK RNAI INDUCED CELLS (48H), SUPPORTING THE DIFFERENTIATION TO PCF PHENOTYPE.....	182
TABLE 5-4: SEQUENCE ORTHOLOGUES OF TbDFK FROM PUBLISHED GENOMES.. ..	193

List of Figures

FIGURE 1-1: THE OVERVIEW OF THE TRANSMISSION CYCLE OF HUMAN AFRICAN TRYPANOSOMIASIS.	16
FIGURE 1-2: DISTRIBUTION OF HUMAN AFRICAN TRYPANOSOMIASIS WITH INCIDENCES AND RISK FOR TRAVELLERS.	19
FIGURE 1-3: MORPHOLOGICAL FORMS OF TRYPANOSOMES OVER THE COURSE OF THE <i>TRYPANOSOMA BRUCEI</i> LIFE CYCLE, DETERMINED FROM ELECTRON MICROGRAPHY..	23
FIGURE 1-4: OVERVIEW OF THE MORPHOLOGICAL STAGES OF <i>T. BRUCEI</i> DEVELOPMENT USING LIGHT AND FLUORESCENCE MICROSCOPY.....	24
FIGURE 1-5: A GRAPHIC DEPICTING THE CELL DIVISION CYCLE OF <i>T. BRUCEI</i> . T	25
FIGURE 1-6: A. STRUCTURE OF PKA WITH SUBDOMAINS MAPPED.	31
FIGURE 1-7: THE TRYPANOSOME KINOME – A STYLISED REPRESENTATION OF THE EUKARYOTIC PROTEIN KINASE COMPLEMENT OF <i>HOMO SAPIENS</i> AND <i>T. BRUCEI</i>	36
FIGURE 3-1: SCHEMATIC DEPICTING GATEWAY CLONING OF INSERTS INTO PGL2084.	79
FIGURE 3-2: PREDICTION OF HAIRPIN FOLDING FROM THE REDESIGNED, 150 NT STUFFER FRAGMENT OF PGL2084..	79
FIGURE 3-3: VALIDATION OF CORRECT CLONING BY RESTRICTION ENDONUCLEASES DIGEST OF PGL2084 AND A CORRECTLY RECOMBINED PLASMID (PTL2).	80
FIGURE 3-4: GROWTH CURVES OF TWO CELL LINES TARGETING CRK3 USING DIFFERENT RNAi SYSTEMS.	81
FIGURE 3-5: GROWTH CURVES FOR SELECTED INDUCED RNAi CELL LINES AND A 2T1 CONTROL, WITH THE CORRESPONDING VALUES OF THE ALAMAR BLUE ASSAY WHICH MEASURES CELL VIABILITY AT 72H.	83
FIGURE 3-6: RESULTS OF A KINOME-WIDE RNAi SCREEN.	96
FIGURE 3-7: BAR CHART SHOWING THE MEAN RELATIVE QUANTIFICATION OF MRNA IN INDUCED AND UNINDUCED RNAi CELL LINES.....	101
FIGURE 4-1: DEPICTION OF THE DOMAIN ARCHITECTURE OF THE PREDICTED NEKS OF <i>T. BRUCEI</i>	114
FIGURE 4-2: THE MECHANISMS OF PDK1 ACTIVATION OF PKB, SGC AND S6K.....	116
FIGURE 4-3: SCHEMATIC REPRESENTATION OF NEK 12.1 PROTEIN DOMAIN ARCHITECTURE WITH KEY FEATURES ANNOTATED.	118
FIGURE 4-4: SEQUENCE ALIGNMENT OF NEK 12.1 AND NEK 12.2 PROTEINS.	119
FIGURE 4-5: GROWTH CURVES FOR <i>T. BRUCEI</i> 2T1 BSF RNAi LINES (15356/15357) UNDERGOING NEK12 DOUBLE KNOCKDOWN.....	122
FIGURE 4-6: RESULTS OF DAPI COUNTS SHOWING PROPORTION OF CELL TYPES IN NEK 12 DEPLETED CELLS	124
FIGURE 4-7: HISTOGRAMS SHOWING THE DNA CONTENT OF NEK 12 RNAi INDUCED AND UNINDUCED CELLS.....	125
FIGURE 4-8: SCHEMATIC DEPICTING THE CONSTRUCT ENCODED TAGS FUSED TO THE C-TERMINUS OF PROTEINS EXPRESSED USING <i>E. COLI</i> IN THIS STUDY.	126
FIGURE 4-9: ANTI-HIS WESTERN BLOT (PANEL A.) SHOWING THE EXPRESSION LEVELS OF VARIOUS NEK AND ORPHAN KINASES.	128
FIGURE 4-10: INHIBITION OF NEK12.1 KINASE ACTIVITY WITH 3-MB-PP1.	131
FIGURE 4-11: HOMOLGY MODELS OF NEK 12.1 (A) AND NEK 12.2 (B) CROSS SECTIONED THROUGH THE ACTIVE SITE TO SHOW THE PREDICTED SIZE OF THE REAR HYDROPHOBIC POCKET.	133
FIGURE 4-12: STRUCTURES OF COMPOUNDS PREDICTED TO DOCK INTO THE NEK 12.1 ACTIVE SITE IDENTIFIED BY VIRTUAL SCREENING.	134
FIGURE 4-13: STRUCTURES OF COMPOUNDS PREDICTED TO DOCK INTO THE NEK 12.2 ACTIVE SITE IDENTIFIED BY VIRTUAL SCREENING	135
FIGURE 4-14: SCHEMATIC REPRESENTATION OF ZFO PROTEIN DOMAIN ARCHITECTURE WITH KEY FEATURES ANNOTATED. THE SCALE ABOVE THE DIAGRAM SHOWS THE AMINO ACID NUMBER.	136
FIGURE 4-15: GROWTH CURVES FOR <i>T. BRUCEI</i> 2T1 BSF RNAi LINES (STL52/55) UNDERGOING ZFO KNOCKDOWN.....	138
FIGURE 4-16: SCHEMATIC REPRESENTATION OF TbPDK1 AND HsPDK1 PROTEIN DOMAIN ARCHITECTURE WITH KEY FEATURES ANNOTATED.	139
FIGURE 4-17: GROWTH CURVES FOR <i>T. BRUCEI</i> 2T1 BSF RNAi LINES (15059/15060) UNDERGOING TbPDK1 KNOCKDOWN	140
FIGURE 5-1: SCHEMATIC REPRESENTATION OF THE DFK PROTEIN	163
FIGURE 5-2: DIAGRAM OF THE POSITION OF PREDICTED TM DOMAINS OF DFK FROM FIVE DIFFERENT PREDICTION ALGORITHMS.	164
FIGURE 5-3: GROWTH CURVES OF DFK RNAi CELL LINES (PLUS/MINUS INDUCTION).....	165
FIGURE 5-4: FACS CURVES FOR CONTROL SAMPLES SHOWING THE FLUORESCENCE OF BSF (A.), PCF (C.) AND MIXED CELLS (B.) WHEN LABELLED WITH A FITC-CONJUGATED ANTI-EP PROCYCLIN MONOCLONAL ANTIBODY.....	167
FIGURE 5-5: FACS CURVES FOR DFK RNAi (CELL LINE STL201) SAMPLES AT 37°C IN HMI-11 SHOWING THE FLUORESCENCE OF RNAi INDUCED CELLS WHEN LABELLED WITH A FITC-CONJUGATED ANTI-EP PROCYCLIN MONOCLONAL ANTIBODY.	168
FIGURE 5-6: A. AN EXAMPLE OF AN APPARENTLY MORPHOLOGICALLY CORRECT PCF AFTER DIFFERENTIATION FROM BSF INDUCED BY RNAi OF DFK FOR 72 H	169

FIGURE 5-7: PANEL A. WESTERN BLOT TO SCREEN PARASITE CLONES TRANSFECTED WITH PGL2203 TO ENDOGENOUSLY TAG DFK WITH C-TERMINAL 12MYC.	172
FIGURE 5-8: IMMUNOFLUORESCENCE LOCALISATION OF DFK-12MYC IN CELL LINE STL2379, USING 2T1 CELLS AS A NEGATIVE CONTROL.	174
FIGURE 5-9: TRANSMISSION IMMUNE-ELECTRON MICROGRAPHS FROM CELL LINE STL2378 TREATED WITH MOUSE ANTI-MYC IGG AND ANTI-MOUSE IMMUNOGOLD (10 NM) SPHERES.	176
FIGURE 5-10: IMMUNOPRECIPITATION OF DFK-12MYC.	178
FIGURE 5-11: TRANSCRIPT EXPRESSION PROFILING OF DFK FROM A. QUIEROZ <i>ET AL.</i> 2009, B. KABANI <i>ET AL.</i> 2009 C. JENSEN <i>ET AL.</i> 2009 AND SIEGEL <i>ET AL.</i> 2010.	188
FIGURE 5-12: SCHEMATIC REPRESENTING THE KNOWN REGULATORS OF BSF TO PCF DIFFERENTIATION. KNOWN INTERACTIONS ARE CONNECTED BY ARROWS.	195

List of Accompanying Material

Supplemental CD-ROM Containing:

S1: List of oligonucleotides, plasmids and cell lines used in kinome-wide RNAi screen.

S2: Molecular models and templates of NEK 12.1 and NEK12.2, results of data mining and virtual screening.

S3: List of differentially regulated genes identified following RNA-seq transcriptomic analysis of DFK RNAi cell lines.

Abbreviations

3-MB-PP1	1-(<i>tert</i> -Butyl)-3-(3-methylbenzyl)-1H-pyrazolo[3,4-d]pyrimidin-4-amine
aa	amino acid
ATP	adenosine triphosphate
BLEO	bleomycin / phleomycin
bp	base pair
BSA	bovine serum albumin
BSD	blasticidin
BSF	bloodstream form <i>T. brucei</i>
CATT	card agglutination test for trypanosomiasis
cDNA	Complementary DNA
CDS	coding sequence
DAPI	4,6-diamidino-2-phenylindole (nucleic acid stain)
DDT	dichlorodiphenyltrichloroethane
DFMO	α -difluoromethylornithine
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DTT	dithiothreitol
EDTA	ethylene diamine tetra acetic acid
EF-1 α	elongation factor 1 a
FACS	fluorescence activated cell sorting
FAZ	flagellar attachment zone
FCS	fetal calf serum
gDNA	genomic DNA
GFP	green fluorescent protein
GPI	glycosylphosphatidylinositol
HA	human influenza hemagglutinin
HAT	Human African Trypanosomiasis
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HIS	histidine
HRP	horseradish peroxidase
HYG	hygromycin B
IPTG	isopropyl- β -D-thiogalactopyranoside
K	kinetoplast
Kb	kilo base
kDa	kilo Dalton
LB	Luria bertani medium
LC MS/MS	liquid chromatograph tandem mass spectrometry
LOF	Loss of Fitness

m	milli / metre
M	molar
mRNA	messenger ribonucleic acid
n	nano
N	nucleus
NECT	eflornithine:nifurtimox combination therapy
NP40	octylphenyl-polyethylene glycol
nt	nucleotide
OD	optical density
OPB	oligopeptidase b
ORF	open reading frame
PBS	phosphate buffered saline
PCF	procyclic form <i>T. brucei</i>
PCR	polymerase chain reaction
pTL	Plasmid- Trypanosome Library
PTS	Peroxisomal Targeting Sequence
qPCR	quantitative PCR
RNA	ribonucleic acid
RNAi	ribonucleic acid interference
rRNA	ribosomal ribonucleic acid
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
STL	Stabilate Trypanosome Library
TEM	transmission electron microscopy
TLO	Trypanosome Library Oligonucleotide
UTR	untranslated region
UV	ultra violet
V	volts
v/v	volume to volume
VSG	variant surface glycoprotein
w/v	weight to volume
WHO	World Health Organisation
λ (λ_{Ex} / λ_{Em})	wavelength (excitation λ / emission λ)
μ	micro

Amino acid abbreviations

A alanine (Ala)
C cysteine (Cys)
D aspartic acid (Asp)
E glutamic acid (Glu)
F phenylalanine (Phe)
G glycine (Gly)
H histidine (His)
I isoleucine (Ile)
K lysine (Lys)
L leucine (Leu)
M methionine (Met)
N asparagine (Asn)
P proline (Pro)
Q glutamine (Gln)
R arginine (Arg)
S serine (Ser)
T threonine (Thr)
V valine (V)
W tryptophan (Trp)
X any amino acid
Y tyrosine (Tyr)

Acknowledgements

Firstly I would like to sincerely thank my primary supervisor Jeremy Mottram for his guidance, support and patience over the past four years. I feel very lucky to have been offered the opportunity to work on this project, to learn new concepts and techniques, and to travel to courses and conferences during my PhD. I thank Malcolm Walkinshaw and members of his group for supervising the parts of my project at the University of Edinburgh. Also, I thank my assessors Cathy Moss and Sylke Müller for their support. I thank SULSA for their financial support and stimulating postgraduate programme; it was a great opportunity to make friends and contacts all over Scotland. Working in the Mottram group has been a pleasure, the fantastic technical support from Elaine Brown, Jim Scott and Alana Hamilton made this project possible, thank you for everything. I also thank Tim Mottram for assisting with the generation of the pTL plasmids.

I have to acknowledge the many great friends I have made in Glasgow, I really enjoyed sharing the good times with you and your support in the bad times. I especially appreciate everyone who helped me out when I broke my leg. Thanks to all my work mates for the stimulating and friendly, and slightly distracting environment! Lots of people made Glasgow a great place to be: Alli, Amy, Andy, Ben, Cat, Corinna, Cris, Daniela, Darren, Dave, Ed, Eileen, Elizabeth, Elmarie, Federica, Federico, Jana, Jane, Janet, Joe, Kerry, Kiran, Raquel, Ryan, Sam, Will (and anyone else I have accidently missed here)- I've really valued your friendship and I hope that this will continue for years to come.

Emily, Fillip, Frøya, Mona and Patricia; thanks for being amazing flatmates, I will really miss living with you in Kelvingrove Street. All the guys from Friday football; thanks for putting up with my comically long list of injuries. Anthony, Emelie, Michael, Raph, and the other chemists; thanks for keeping me sane with trips to super circuits, climbing walls, mountains and pubs. Julie; thanks for being a great friend, I'm really glad you came to Glasgow.

I finally thank my family for all their love and support over the years. Mum and Dad, I will simply never be able to thank you enough for all the opportunities you gave me.

“Glasgow is a magnificent city”, said McAlpin. “Why do we hardly ever notice that?”

“Because nobody imagines living here,” said Thaw.

Alasdair Gray, from ‘Lanark’.

1 General introduction

1.1 African Trypanosomiases caused by *Trypanosoma brucei* spp.

The *Trypanosoma brucei* species complex contains three species of morphologically indistinguishable, unicellular protozoa that are responsible for a large number of deaths of humans and livestock across sub-Saharan Africa. These organisms have fascinating, complex biological features and life cycles that belie the suffering they continue to impose on those affected. *Trypanosoma brucei brucei* is a pathogen of cattle (causing the disease nagana), *T. b. rhodesiense* is an incidental, zoonotic pathogen of humans and *T. b. gambiense* is the major trypanosomatid pathogen of humans in Africa - all are transmitted by species of tsetse flies in the genus *Glossina* (Figure 1-1). The human disease African Sleeping Sickness or Human African Trypanosomiasis (HAT), has historically varied in the scale and severity of outbreaks but due to current control techniques is, tentatively, on a trajectory for elimination.

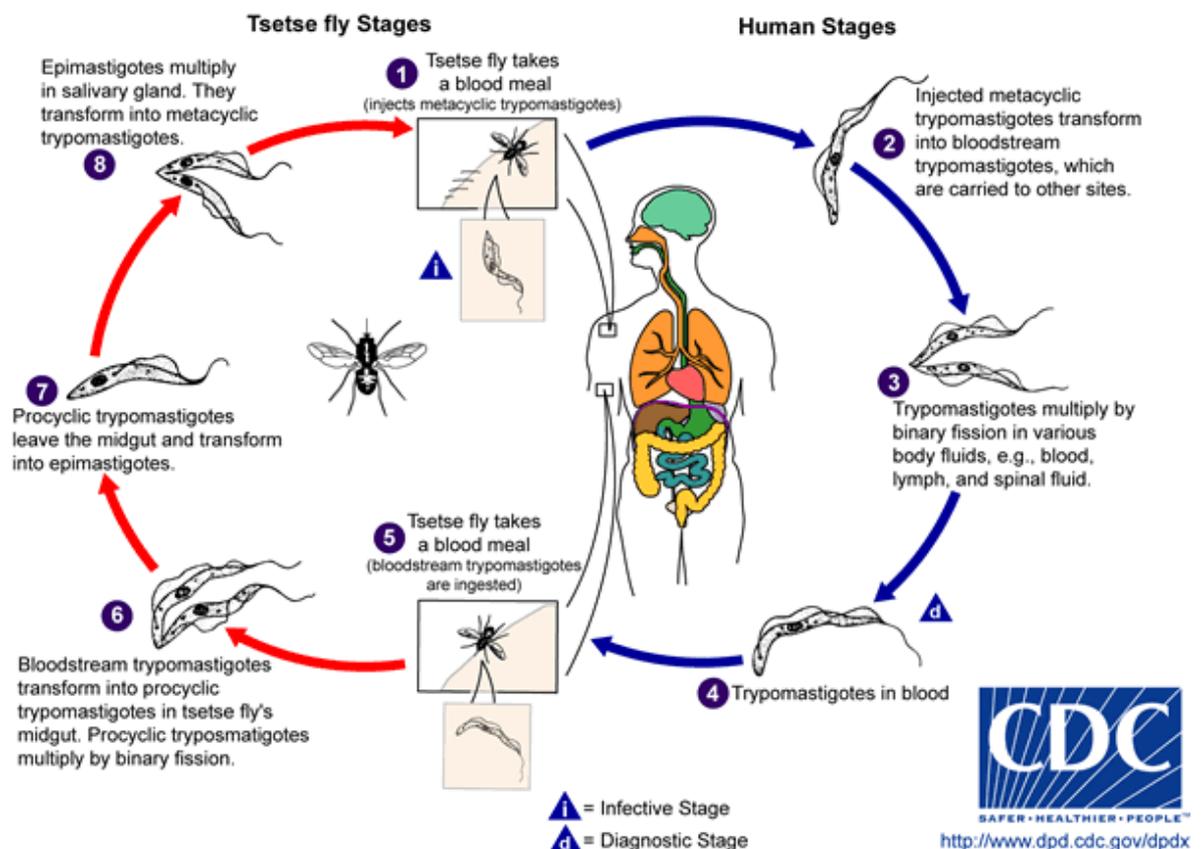


Figure 1-1: The overview of the transmission cycle of Human African Trypanosomiasis.

Before discussing the biology of *Trypanosoma brucei* the problem of HAT should be reviewed for context.

HAT has in the past afflicted vast swathes of the African continent and currently anyone who lives within the “tsetse belt”, between 14°N and 20°S, is at risk of infection (Malvy and Chappuis, 2011). In practice, however, the disease has tended to be associated with distinct foci usually based on the ecology of the tsetse vector species and where humans contact this range. Currently the incidence of HAT cases is at an historic low with only 7139 reported cases in 2010 (WHO, 2012 - www.who.int/trypanosomiasis_african/en/). A number of factors mean this is undoubtedly an underestimate. Despite this, active case finding, drug treatment and integrated vector control are controlling the disease. The two forms of human disease are geographically segregated (with one exception)- the chronic disease caused by *T. b. gambiense* is confined to West and Central Africa and the acute disease caused by *T. b. rhodesiense* is found in East Africa. Only in Uganda are the two forms found in the same geographic area (Figure 1-2).

The history and prehistory of these parasites and disease are reviewed well by Steverding, 2008. African trypanosomiasis are ancient diseases, with salivarian trypanosomes predicted to have evolved 100-300 million years ago with tsetse flies predicted to be around 35 million years old. It is likely that early hominids were exposed to these parasites and this had an effect on human evolution, selecting for trypanotolerance -only two of the African trypanosomes now infect humans. This is supported by the fact that local grazing animals are tolerant to trypanosomes and domesticated livestock are susceptible to trypanosomiasis (Lambrecht, 1985). Reports of diseases resembling nagana and HAT have been proposed to exist in records from the ancient Egyptians and the Middle Ages. The disease was first documented in early modern Western reports in the 1700's due to investigations of medical officers in slave-trading companies but it was not until the 1800s that the association between tsetse flies, trypanosomes and (animal) trypanosomiasis was determined. Rigorous investigations into this disease were carried out by David Bruce, who in 1895 demonstrated that trypanosomes were the causative agent of nagana. Further investigations by Robert Forde (1901) and Joseph Dutton (1902) identified trypanosomes in the blood of humans; trypanosomes were isolated from cerebrospinal fluid (CSF) by

Aldo Castellani. Eventually Bruce was able to demonstrate the complete cyclical life cycle of *T. brucei* through tsetse flies.

The past century has seen three large-scale epidemics of HAT - the first, between 1896 and 1906 killed almost 1 million people and led to major efforts to develop therapies. The early treatments were toxic arsenicals (atoxyl, trypanemide) and it was only in 1916 that an effective and (for that time) relatively safe drug, suramin, was developed. Incredibly, almost 100 years later, this drug is still in use. Between 1920 to 1940 another epidemic raged, leading to control by active case finding with drug treatment and vector/reservoir control. During this time both pentamidine and DDT were developed which assisted with control methods. The drug melarsoprol was discovered in 1949 and became and the only treatment for the second stage of *T. b. rhodesiense*. By the 1960s trypanosomiasis was under good control. Between the mid-1960s until the early 1990s the number of cases rose due to a number of reasons. As African nations gained independence from colonial powers many descended into political and economic instability and a number experienced internal conflict or war. As HAT was at a low prevalence it became a low priority for health services in many countries. These factors led to a resurgence of cases across Africa, but with particularly in countries experiencing war or civil unrest - which also hampered efforts to control the epidemic. A new drug did become available, in the form of eflornithine (or DFMO), which was much safer than melarsoprol for treating second stage *T. b. gambiense* disease. A renewed emphasis on HAT control in the 1990s to 21st century has led to the number of confirmed cases dropping below 10, 000. With further intense efforts it may be possible to eliminate *T. b. gambiense* disease from large areas or even totally if the status quo remains. However, experience shows that if attention is not paid to controlling the disease, or if circumstances do not permit effective control, then outbreaks can flare rapidly and dramatically (Steverding, 2008).

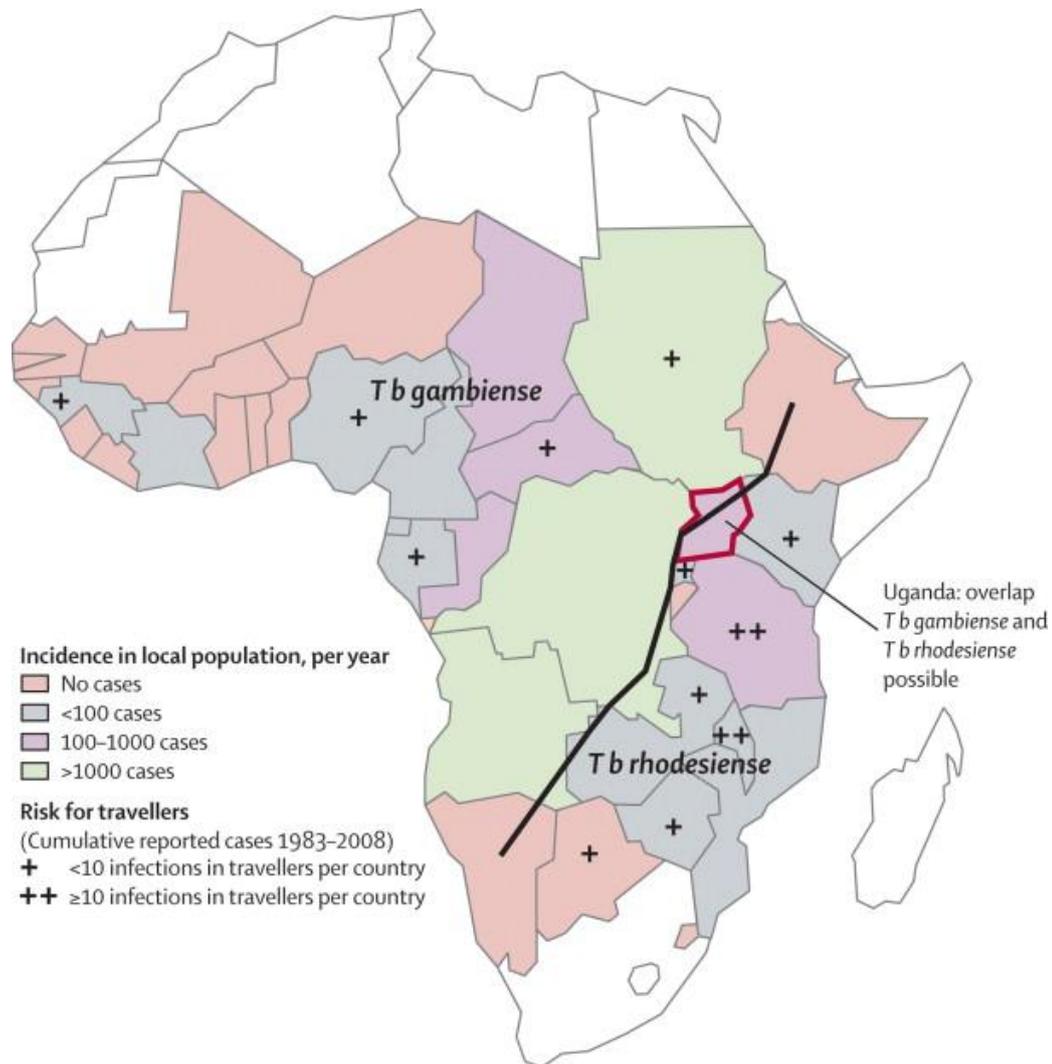


Figure 1-2: Distribution of human African trypanosomiasis with incidences and risk for travellers. The black line divides the areas in which *Trypanosoma brucei gambiense* prevails and those in which *Trypanosoma brucei rhodesiense* predominates – Reproduced with permission from (Brun et al., 2010)

Infection with *T. b. gambiense* is generally fatal if untreated - until recently much of the literature stated it was invariably fatal. However, evidence from a long term follow up study of patients refusing treatment for HAT appears to show evidence for trypanotolerance and self cure (Jamonneau et al., 2012). These findings have implications for control efforts (i.e. detection of silent carriers) and statisticians (estimating the amount of under-reporting of cases). For most patients though, disease is severe. The typical clinical progression recently reviewed in (Brun et al., 2010; Malvy and Chappuis, 2011) states that following infection a localised chancre can form, though rarely in *T. b. gambiense* cases and more often in *T. b. rhodesiense* cases, this can be followed by lymphadenopathy or Winterbottom's sign. The parasites then typically enter the bloodstream and replicate, parasitaemia waves - correlated

with VSG switching - leading to waves of fever and malaise. This is the haemolymphatic stage (or Stage 1) of the disease. Symptoms are variable with *T. b. gambiense* infection but the general trend is chronic and progressive - with typically 1.5 years of Stage 1 disease before entry into Stage 2, the meningoencephalitic stage. This stage typically also persists for 1.5 years and results from parasites actively penetrating the blood-brain barrier. Here the disease exhibits the symptoms for which it is named; major disruption of the circadian rhythm can occur, leading to disturbed sleep/wake patterns. Other damage can lead to generalised neurological indications including weakness, paralysis and movement disorders - a broad range of psychiatric disorders can arise. Death typically occurs within 4 years of infection with *T. b. gambiense*. The clinical picture for *T. b. rhodesiense* is much more severe and acute, with disease leading to death in weeks to months, though this only accounts for around 2% of HAT cases.

Due to the non-specific symptoms of Stage 1 disease, diagnosis for *T. b. gambiense* is carried out using the card agglutination test for trypanosomiasis (CATT) - this test is practical to use in resource poor settings, fast and relatively robust. CATT allows for mass screening to be performed using fingerprick blood but a positive result needs to be confirmed by microscopy. The stage of the disease must also be determined, so a parasitaemic patient must then have a lumbar puncture followed by assessment of CSF for parasites or high levels of lymphocytes. Tests may have to be repeated over time to ensure accurate diagnosis.

Treatment of the disease varies with the species and stage of the disease. Four main treatments are currently in use. First stage *T. b. gambiense* is treated with pentamidine and the second stage is now treated with eflornithine:nifurtimox combination therapy (NECT) - alternatives to this are eflornithine monotherapy and melarsoprol. *T. b. rhodesiense* infection can be treated in the first stage with suramin and second stage with melarsoprol. Pentamidine is generally well tolerated but requires daily dosing for a week by intramuscular injections - it is only useful against Stage 1 *T. b. gambiense* disease. Suramin is effective against *T. b. gambiense* but due to the risk of adverse reactions in patients co-infected with *Onchocerca* spp. is rarely used in West Africa. It does generate frequent adverse reactions and the regimen is 30 days long. Melarsoprol is a poorly

tolerated drug and it is commonly presented as the proverbial “Catch-22”; Stage 2 disease must be treated or else prove fatal but between 4-8% of patients treated for stage 2 disease with melarsoprol will develop an encephelopathic syndrome leading to death in 44-57% of instances. Treatment with eflornithine monotherapy is more efficacious than melarsoprol and results in far fewer adverse effects however it requires a dosing regimen of four infusions per day over two weeks. Dosing kits have been developed by WHO but this is still a complex drug to administer in resource-poor settings (Brun et al., 2010). In order to reduce the length of time of treatment a combination therapy of eflornithine and nifurtimox was developed. This 7 day regime combines eflornithine infusions with oral nifurtimox, a drug originally used for the treatment of Chagas Disease (caused by *T. cruzi*). This is now being used as the main treatment against *T. b. gambiense* second stage disease in the countries containing 96% of disease cases. More recently Fexinidazole, a new nitroimidazole drug which is orally available and can treat both stages of both diseases in mouse models, has passed phase one clinical trials and is entering a phase II/III trial in late 2012 (Torreele et al., 2010; Mäser et al., 2012; Simarro et al., 2012).

Control of the parasite’s vector is another aspect of successful control strategies. *T. brucei* spp. are transmitted by tsetse flies (both male and female), of which there are 3 groups of species - those that inhabit savannah, forest or riverine environments. They exhibit differing propensities for human feeding and as such vary in their vectoral capacity, though the female flies can feed ever 2-3 days. The locale that tsetse flies inhabit can provide the foci for transmission of HAT, for example watering places, river fords and lakesides. Tsetse flies spend most of the day (23 hours) at rest in vegetation, and avoid temperatures over 36°C. Salivary gland infection rates rarely exceed 0.1%, even in endemic areas. The rate of reproduction for tsetse flies is generally low due to their larviparous life cycle, where the female fly develops a single larva in her uterus every 9-12 days. Tsetse flies are still exceptionally sensitive to all classes of insecticide. These factors combined mean that effective control of tsetse flies by trapping, ultra-low-volume spraying and (in one example) sterile insect technique is possible, and can have dramatic effects on the incidence of HAT (Service, 2004).

1.2 *Trypanosoma brucei* life history, cellular and molecular biology

Trypanosoma brucei species fall into the Eukaryotic supergroup Excavata, group Discoba (Euglenozoa), subgroup Kinetoplastidae, family Trypanosomatidae (Reviewed in (Walker et al., 2011a)). Kinetoplastids are characterised by having a single, large mitochondrion containing its mtDNA as a single body, the kinetoplast. Trypanosomatids possess a single flagellum and are usually parasitic as opposed to Bodonidae (the other order contained within Kinetoplastidae), which typically possess two flagella and are usually free living. They are considered early branching eukaryotes and have become some of the best categorised protozoa due to their amenability for genetic manipulation and *in vitro* culture.

African trypanosomes have a complex life cycle involving a mammalian host and a tsetse fly vector (Figure 1-3 and Figure 1-4). They encounter varied environments during their life history and undergo spectacular morphological changes accompanied by transcriptomic, proteomic and metabolic changes. They exist in the mammalian host as two morphological forms, a replicative, slender trypomastigote and a quiescent stumpy trypomastigote and are often termed the bloodstream forms (BSF). The stumpy form is preadapted for transmission to a tsetse fly (discussed in detail in Chapter 5). Once ingested by a tsetse fly the parasites encounter a harsh environment in the digestive tract of the flies and as such transform into a procyclic trypomastigote form able to survive there. Procyclic forms (PCF) multiply in the midgut and transform into mesocyclic trypomastigotes, which migrate to the foregut. Here they transform into epimastigotes that divide asymmetrically to form a long and short epimastigote. They then migrate to the salivary glands where the short epimastigotes can attach to the lining of the salivary gland using their flagellum to form a flaglipodium. The epimastigotes can divide asymmetrically to form new epimastigotes and infective metacyclic trypanosomes which are injected into a new mammalian host during the tsetse bloodfeeding. This allows tsetse flies to remain infective over their relatively long lifespan (Rotureau et al., 2011, 2012).

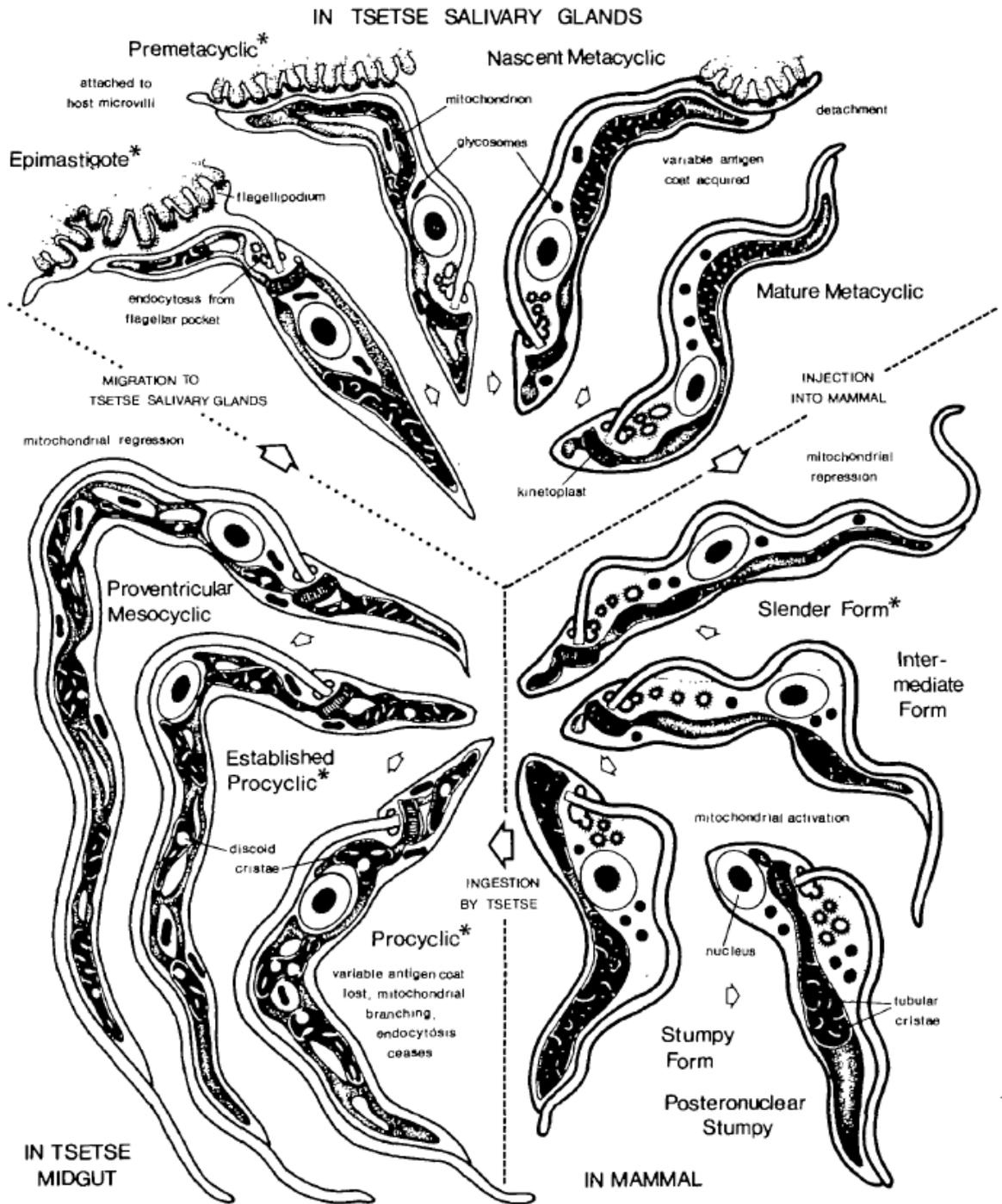


Figure 1-3: Morphological forms of trypanosomes over the course of the *Trypanosoma brucei* life cycle, determined from electron micrography. Reproduced from (Vickerman, 1985) with Permission.

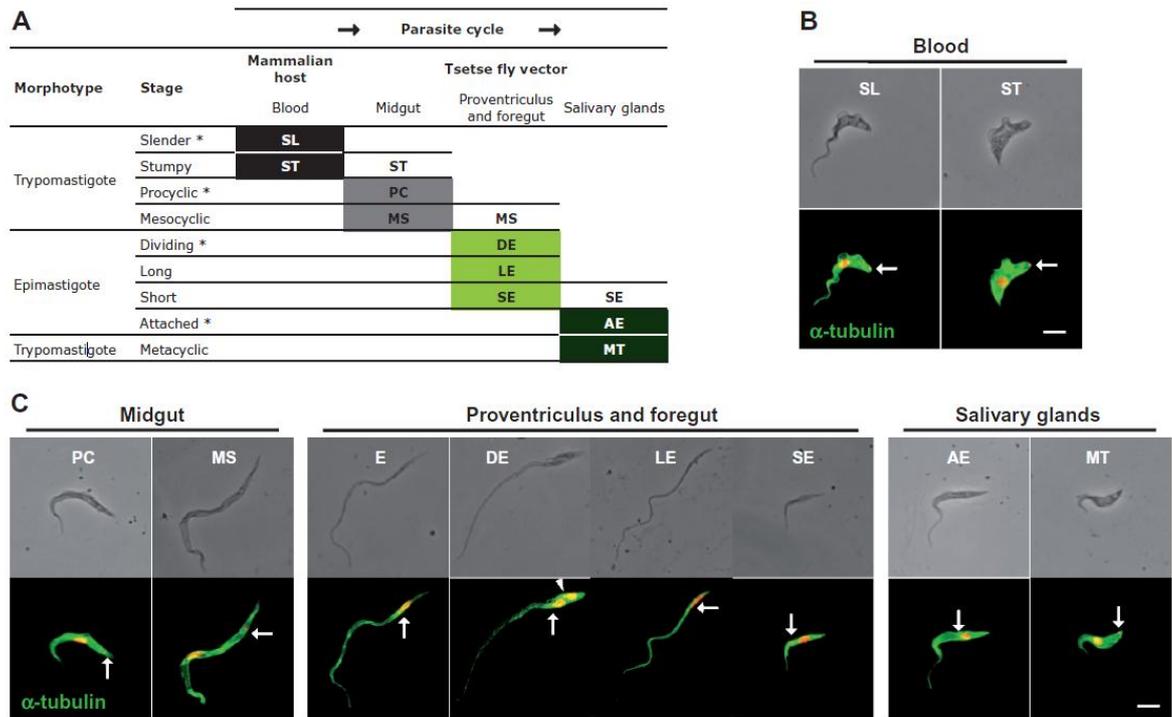


Figure 1-4: Overview of the morphological stages of *T. brucei* development using light and fluorescence microscopy (labelling the cytoskeleton) – reproduced from (Rotureau et al., 2011). Panel a. shows a table listing the localisation of the life cycle stages throughout the life cycle. Panel b. shows the forms found in the mammalian host (SL: slender trypomastigote, ST: stumpy trypomastigote). Panel c. shows the forms found in the tsetse fly. (PC: procyclic trypomastigote, MS: mesocyclic trypomastigote, E: proventricular epimastigote, DE: asymmetrically dividing epimastigote, LE: long epimastigote, SE: short epimastigote, AE: attached epimastigote, MT: metacyclic trypomastigote). * Denotes dividing forms of the parasite.

The forms of *T. brucei* amenable to tissue culture are the PCF and BSF. These are trypomastigote form parasites, with the kinetoplast posterior to the nucleus. In some respects they are an ideal model organism for cell biology due to their single-copy-organelle layout (for most organelles). The cells are typically elongated in shape and have a highly polarised structure. This structure is defined by a sub-pellicular microtubule cytoskeleton (corset), with all the microtubules running in an anterior to posterior direction (i.e. the positive end is posterior) apart from a specific microtubule quartet which runs in the opposite direction. The microtubule corset has a posterior opening to accommodate the flagellar pocket and a gap to accommodate the flagellar attachment zone (FAZ) filament running beside the microtubule quartet. The flagellum, containing a 9+2 axoneme and a paraflagellar rod exits the cell from the posterior flagellar pocket. The axoneme of the flagellum associates with the basal body which is also attached to the kinetoplast via the tripartite attachment complex. The flagellar pocket is the sole point of endo- and exocytosis and is held closely

around the flagellum by a collar structure (Matthews, 2005; Farr and Gull, 2012). During the cell cycle the basal body duplicates and segregates and this is coupled to the division of the kinetoplast. S-Phase in trypanosomes consists of first duplicating and segregating the kinetoplast to achieve a 1N2K configuration and then replicating and the nucleus. The positioning of these organelles is different between the BSF and PCF life cycle stages (Figure 1-5). In the precytokinetic G2 phase, 1N2K cell undergo mitosis and become 2N2K cells which are in M Phase. During S-Phase a new flagellum is grown with its anterior end attached to the old flagellum. This progresses along the old flagellum until the distal, anterior end is reached. Near this point a cleavage plane for cytokinesis is initiated, this furrow ingresses until the two daughter cells are attached at their posterior ends, which then detach in the process known as abscission (Farr and Gull, 2012; Li, 2012).

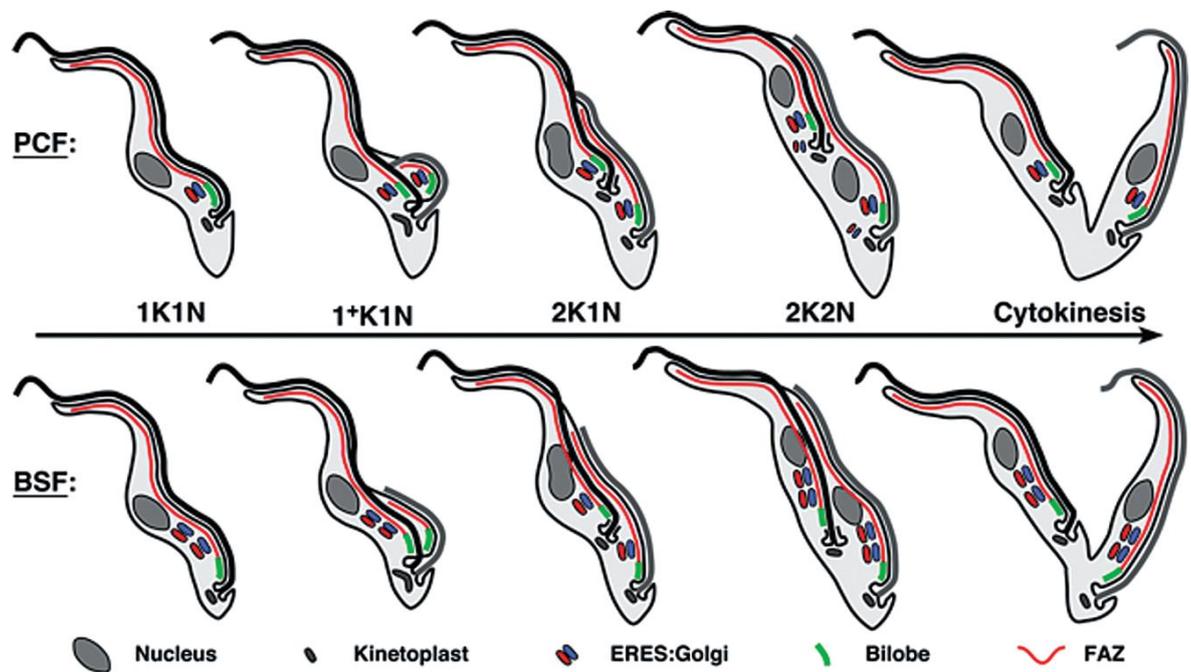


Figure 1-5: A graphic depicting the cell division cycle of *T. brucei*. The lower panel indicating BSF. Reproduced from (Bangs, 2011), with permission.

The *T. b. brucei* TREU 927 genome strain contains a nuclear and a kinetoplast genome totalling 35 megabases (Mb) per haploid genome (Berriman et al., 2005). The nucleus contains 11 pairs of megabase chromosomes (0.9-5.7 Mb), 2-4 intermediate size chromosomes (300-900 kb) and approximately 150 minichromosomes (50-100 kb) that mainly contain silent VSGs. There is a 500 Mb segmental duplication between chromosome 4 and 8 leading to 75 gene

duplications unique to *T. b. brucei* (Jackson, 2007). The total number of genes predicted to be encoded in the 11 megabase chromosomes was 9068 (Berriman et al., 2005). This has been expanded to 11425 following transcriptomic analyses of *T. brucei* which identified novel transcripts (TriTrypDB) and (Siegel et al., 2010). *T. brucei* is predominantly diploid, though meiotic events do occur indicating a haploid life stage at some point in the tsetse fly (Peacock et al., 2009, 2011). The subtelomeric regions contain expression sites for VSG genes (Barry et al., 2005, 2012; Berriman et al., 2005). The mitochondrial genome (kinetoplast) consists of a complex, concatenated network of DNA consisting of minicircles and maxicircles (Liu et al., 2005). The maxicircles encode mitochondrial genes, although once transcribed these require modification by RNA editing to form complete mRNAs. The guide RNAs required for this process are encoded in the minicircles. RNA editing was first described in trypanosomes but is now known to occur in other organisms reviewed in (Simpson et al., 2004). Loss of the mitochondrial genome can be induced in *T. brucei* - mimicking natural dyskinetoplastic parasites such as *T. evansi* and *T. equiperdum* (Schnauffer et al., 2002).

Most genes encoded in the megabase chromosomes are transcribed by RNA polymerase II, in polycistrons with apparently little or no transcriptional control (Palenchar and Bellofatto, 2006; Günzl, 2010; Rudenko, 2010; Preußner et al., 2012). Though transcription is polycistronic, genes involved in the same biological pathways are not arranged into operons. Transcription initiates bidirectionally from strand switch regions - where polycistronic units diverge from each other. The pre-mRNAs are then *trans*-spliced to a 39 nt spliced leader RNA which provides the 5' cap and ribosome binding site - polyadenylation is also performed to mature the mRNA (Ullu et al., 1993). Transcription initiation sites appear to be regulated by epigenetic modifications such as histone type and positioning. Not all of the protein-coding genome is transcribed by RNA Pol II; the active variant surface glycoprotein (VSG) expression site is transcribed by RNA Pol I, in addition pre-rRNAs being generated by this enzyme.

The VSG coat is a distinctive feature of trypanosomes and forms a defence against the antibody driven immune response at cellular and population level. It is transcribed from a single, sub-telomeric expression site - and only a single VSG variant is expressed at a time. This protein is a homodimer, anchored into the

plasma membrane by covalent linkage to glycosylphosphatidylinositol (GPI) and forms a dense monolayer over the surface of the trypanosome. VSG composes 95% of the external cell surface proteins of BSF trypanosomes - this equates to 15% of total cellular protein and $\sim 5 \times 10^6$ VSG dimers (Jackson et al., 1985; Schwede and Carrington, 2010). During an infection, the host mounts an immune response against the predominant VSG variant, however, the antibodies can be cleared by the trypanosomes. This occurs by hydrodynamic forces pushing IgG- and IgM-VSG complexes to the posterior of the cell, where they can be internalised by endocytosis at the flagellar pocket, antibody is removed and degraded in the lysosome. The rate of this is higher in stumpy form parasites. (Engstler et al., 2007). VSG is recycled back to the plasma membrane and trafficked for degradation in the lysosome. This strategy suffices in the short term but is not a suitable tactic over the long course of an infection (2-3 years). Eventually the antibody titre against a specific VSG variant will rise so high that the parasites will be overcome. In order to ensure the survival of the population parasites will switch the active VSG either by switching the active expression site, or by switching the VSG gene transcribed from an expression site. This is achieved by homologous recombination targeted to double-stranded breaks occurring in 70 bp repeats occurring upstream of an active VSG. VSG genes can be recombined from a massive library (~ 2000) of complete VSGs and pseudogenes which can be combined into fully functional mosaic VSGs (Rudenko, 2011; Barry et al., 2012).

One of the features unique to trypanosomes (over the mammalian host) is the compartmentalisation of glycolysis in BSF parasites into organelles called glycosomes. BSF trypanosomes are reliant solely on the constant supply of glucose from the host for generating ATP by glycolysis to meet its free-energy needs. Glycosomes are single membrane vesicles, similar to peroxisomes. They contain nine glycolytic enzymes, 7 of which convert glucose to 3-phosphoglycerate. This then enters the cytosol for conversion to pyruvate over three steps yielding 1 molecule of ATP. The high rate of glycolysis sustained by these parasites and the lack of regulation of the activity of hexokinase and phosphofructokinase creates a “turbo design” glycolysis. If not compartmentalised this becomes toxic to the cell. Glycosomes are turned over during differentiation between lifecycle stages. Proteins can be targeted to

glycosomes using one of two targeting sequences PTS1 or PTS 2. Due to the differences between the regulation of glycolysis in trypanosomes and humans, and their importance in linking many metabolic pathways they have been examined as drug targets (Michels et al., 2006; Coley et al., 2011; Gualdrón-López et al., 2012).

1.3 Protein Kinases

Protein kinases are enzymes that modify other proteins by the post-translational chemical addition of the γ -phosphate group of ATP. They play regulatory roles in almost all cellular process and when disregulated can cause pathologies. They have eventually become the targets of many drug development programmes and a number of protein kinase inhibitors are now in clinical use.

Protein kinase activity was first described in 1954 (Burnett and Kennedy, 1954), though processes regulated by it had been studied in the 1930s (without the investigators understanding). Subsequently the phosphorylation of glycogen phosphorylase by phosphorylase kinase was demonstrate by Fischer and Krebs, who showed it was dependent on Mg-ATP and that the γ -phosphate group of ATP was transferred to a specific serine residue on phosphorylase (Fischer and Krebs, 1955; Krebs and Fischer, 1956; Fischer et al., 1959). This discovery won the authors a Nobel Prize. Since 1954 the literature has expanded enormously and the importance of protein kinases in widespread biological processes has become apparent - a search for “protein kinase” in PubMed returns over 41, 000 results. A history of key discoveries is provided by Cohen (Cohen, 2002).

Protein kinases in eukaryotes are almost all contained within the eukaryotic protein kinase (ePK) superfamily with a smaller division of atypical protein kinases (aPK). The major divisions of ePKs are along substrate specificity at the highest order, into tyrosine kinases and serine/threonine kinases (which are ubiquitous in eukaryotes). After molecular cloning and sequencing of a number of protein kinases, an analysis of the conserved sequences allowed the catalytic domain and subdomains to be defined (Hanks et al., 1988). This analysis categorised known serine, threonine and tyrosine kinases to determine the core features and key residues of protein kinases. The ePKs are defined by a catalytic region with 12 subdomains. Knowledge of the conserved sequences of protein

kinases combined with genome sequencing has led to the ability to predict the protein kinase complement of a given genome. This complement is termed the kinome. ePKs are among the most common protein families, composing 1.5-2.5% of an organism's genome (Manning et al., 2002a). Using this type of analysis the human genome was predicted to contain over 500 protein kinases, which were classified into various families (Manning et al., 2002b). Protein kinases have been classified into a number of families, based on the sequence similarities of the catalytic domain - 6 major families exist: AGC, CAMK, CMGC, TK, TKL, STE (definitions of acronyms in Section 1.5). A category of "Other" ePKs, falling outside of these families, exists. Proteins that have a different domain structure to ePKs, but still act as protein kinases are termed atypical protein kinases (aPKs) and contain groups such as the PI3-Kinases. The different families share substrate preferences and can have similar modes of regulation or function. For example, the CMGC family contain many cyclin dependent kinases (CDKs), which require the binding of a cyclin partner to become activated. Other CMGCs include the mitogen activated protein kinases (MAPKs), which require activation by phosphorylation of a TxY motif in the activation loop. Some CMGCs have dual specificity for tyrosine as well as serine/threonine (DYRKs). The activators of the MAPKs, which often exist in cascades, are contained in the STE group and form a hierarchy (e.g. MAPKK, MAPKKK, MAPKKKK). The TK (tyrosine kinases) are often receptor kinases and are mostly found in metazoans, mammalian receptor kinases are almost all tyrosine kinases. AGC kinases contain several members that are activated by cyclic nucleotides; CAMKs are regulated by calcium binding (Hanks and Hunter, 1995). Techniques to determine the function and interactions of these proteins are now highly advanced for conventional model organisms and mammalian cells and tissue and high level of understanding exists for a great many protein kinases (Johnson and Hunter, 2005).

The X-ray crystal structure of a protein kinase was first determined for PKA (Protein Kinase A), a cyclic-AMP dependent protein kinase, which began to allow the understanding of structure: function relationships for these enzymes (Knighton et al., 1991). The structure is bilobed, with a cleft between the lobes that is responsible for binding MgATP and the peptide substrate. The conserved amino acid sequences could then be mapped to a three dimensional structure helping to shed light on their importance in protein kinase activity, these are

represented in Figure 1-6 and reviewed in (Hanks et al., 1988; Hanks and Hunter, 1995; Adams, 2001). The N-terminal lobe consists mostly of β -sheets and contains domains I-IV, which are involved in binding and orientating the MgATP. The C-terminal lobe is larger, consists of mainly α -helices, contains domains VIA-XI and is primarily responsible for binding the peptide substrate and catalysing the phosphotransfer reaction. Subdomain V connects the two lobes and acts as a hinge region. There are a number of invariant or almost invariant residues and motifs spread throughout the subdomains of the kinase domains. These are indicated in Figure 1-6b, most are involved in binding ATP into the active site and/or catalysing the phosphotransfer reaction. The GxGxxG motif of subdomain I interact with the α - and β - phosphates of the ATP and the invariant lysine of subdomain II facilitates the transfer of the phosphoryl group. The lysine of subdomain II is almost always required for protein kinase activity. The glutamic acid (E) of subdomain III also stabilises the interaction of the α - and β - phosphates with the invariant lysine. The next conserved motif is DxxxxN occurring in subdomain VIB; this is the catalytic loop. The aspartate is extremely important for catalytic activity and asparagine helps stabilise the loop and also binds magnesium in the active site. Subdomain VII contains the highly conserved DFG motif which stabilises another Mg^{2+} , which in turn orients the β - and γ - phosphates into the active site - its orientation is crucial for kinase activity. The APE motif in subdomain VIII provides stability within the C-terminal lobe by the glutamic acid, forming an ion pair with an invariant arginine in subdomain XI. Subdomain VIII also contains residues important in peptide substrate recognition and contains the activation loop found in many protein kinases, which often requires threonine (and a tyrosine in MAPKs) to be phosphorylated to activate the protein kinase. Subdomain IX contains a conserved DxxxxG motif that plays a role in stabilising the catalytic loop and peptide binding domain. Subdomain X is the least conserved region. It is this conservation in the structural features and function of protein kinases that led them to be labelled as “undruggable” for many years, as blocking ATP binding in one kinase using an inhibitor would do it for all kinases and thus be toxic (Nakano et al., 1987; Fedorov et al., 2010).

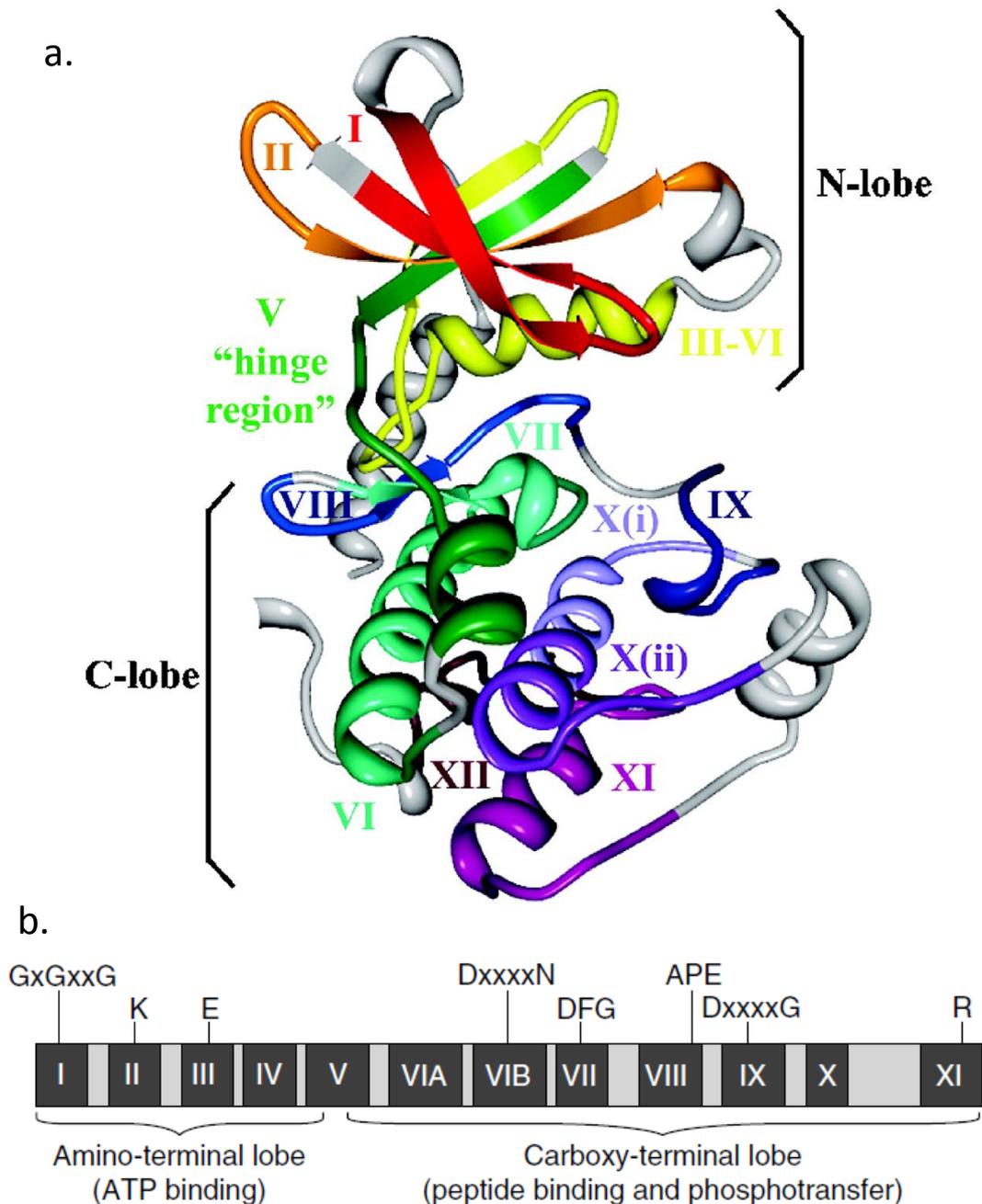


Figure 1-6: a. Structure of PKA with subdomains mapped – Adapted from (Torkamani and Schork, 2008) and b. linear representation of the subdomains with invariant motifs annotated (Hanks, 2003).

1.4 Protein Kinases as Drug Targets

Even until the late 1990's the prevailing dogma was that developing a selective protein kinase inhibitor would be impossible or very difficult. This dogma was broken by the development of Gleevec (Imatinib) for treating chronic myeloid leukaemia (CML) during this time researchers at Ciba-Geigy (Novartis) identified compounds with activity against the BCR-ABL kinase, reviewed in (Deininger and Druker, 2003; Druker, 2004). This is a product of a chimeric gene caused by a

chromosome translocation - the protein possesses constitutive tyrosine kinase activity and activates various signalling pathways that generate a pathological increase in haematopoietic stem cells and myeloid cells. Through high throughput screening and analysis of structure:activity relationships the team identified a class of 2-phenylaminopyrimidine compounds active against platelet-derived growth factor receptor (another receptor tyrosine kinase). After various modifications to improve potency and selectivity they were also discovered to possess activity against BCR-ABL kinase. Further modifications to increase solubility generated lead compounds with Imatinib being progressed into Phase I clinical trials in 1998. *In vitro* this could inhibit BCR-ABL with a IC₅₀ of 250 nM. It proved to be highly effective against CML in patients, even those who had the most severe stage of the disease, and also those resistant to the other therapies then in use. It was also extremely well tolerated by patients, and apparently had a wide therapeutic window and high bioavailability from oral dosing. Phase II and III trials further showed the efficacy of imatinib over the existing treatments. Imatinib was once thought to have exquisite sensitivity, due to it binding the inactive form of the BCR-ABL kinase and by contacting 21 amino acids. It is a Type II protein kinase inhibitor (see below), this also contributes to the low IC₅₀ as it is not an ATP competitor (Type I inhibitor). This success story broke the dogma of protein kinases not being suitable as drug targets. It also provides an example of target based drug development - a one gene, one protein one inhibitor approach - that was very prevalent during the 1990s and 2000s.

Due to the clinical success of Imatinib, the demonstration of the possibility to achieve selective inhibition of a protein kinase and the increases in the understanding of protein kinase involvement in various pathologies there has been an explosion in the level of research in this area. There are over 100 protein kinase inhibitors in either clinical use or trials and many more as preclinical candidates (Barouch-Bentov and Sauer, 2011). The range of diseases that involve protein kinase dysfunction has now expanded to include various cancers, inflammatory/immunological disorders and neurodegenerative diseases (Rokosz et al., 2008; Cuny, 2009; Zhang et al., 2009).

The ability to generate inhibitors specific to a given protein kinase *within* a species has led researchers to reason that selective inhibition of a protein kinase in a *different* species is possible and will allow this strategy to generate new

treatments for use against infectious diseases. Currently attempts to achieve this are underway for trypanosomatid parasites and apicomplexan parasites (Doerig et al., 2005; Naula et al., 2005; Brumlik et al., 2011). It has now become acknowledged that the specificity of protein kinase inhibitors can vary greatly, ranging from the most “selective” ones such as lapatinib (inhibiting a couple of non-targets protein kinases) to staurosporine- which inhibits virtually every protein kinase (Karaman et al., 2008). Similarly the strategy of target based drug development is also under question - due to the complex nature of many diseases. The term polypharmacology (or network pharmacology) has come into parlance, and the emphasis with this field is to target multiple proteins in a pathway using multiple inhibitors in combination or to use a single inhibitor to inhibit multiple targets (Hopkins, 2008). It may be that protein kinase inhibitors are ideal for the latter role (Boran and Iyengar, 2010; Knight et al., 2010).

Protein kinase inhibitors can typically be broken down into three major types. Type I inhibitors are typically ATP competitive and bind to a protein kinase in its active conformation. Type II inhibitors bind a protein kinase in a region adjacent to the ATP binding site, which is exposed when in its inactive conformation, and therefore are not competitive with ATP. This conformation is called “DFG out” due to the conformation of this conserved motif. These types of inhibitor lock the protein kinase in an inactive conformation. Type III and Type IV inhibitors are those that bind away from the active site but cause allosteric changes that inactivate the protein kinase or prevent it from being activated. Type III inhibitors tend to bind closer to the active site than Type IV inhibitors (Knight and Shokat, 2005; Liu and Gray, 2006). Inhibitors that do not compete with ATP are typically more potent than ATP competitive inhibitors due to the high intracellular ATP concentration of 1-10 mM.

1.5 Protein Kinase of *T. brucei*

In 2005 the release of the genomes of three pathogenic trypanosomatids (*T. brucei*, *T. cruzi*, *Leishmania major*) allowed for bioinformatics searches to examine them for a comprehensive prediction of their protein kinase complement (Ivens et al., 2005; Parsons et al., 2005a). This revealed that *T. brucei* had a predicted complement of 156 ePKs and 20 aPKs (as well as 13 ePK proteins predicted to be inactive- pseudokinases). These were then compared to

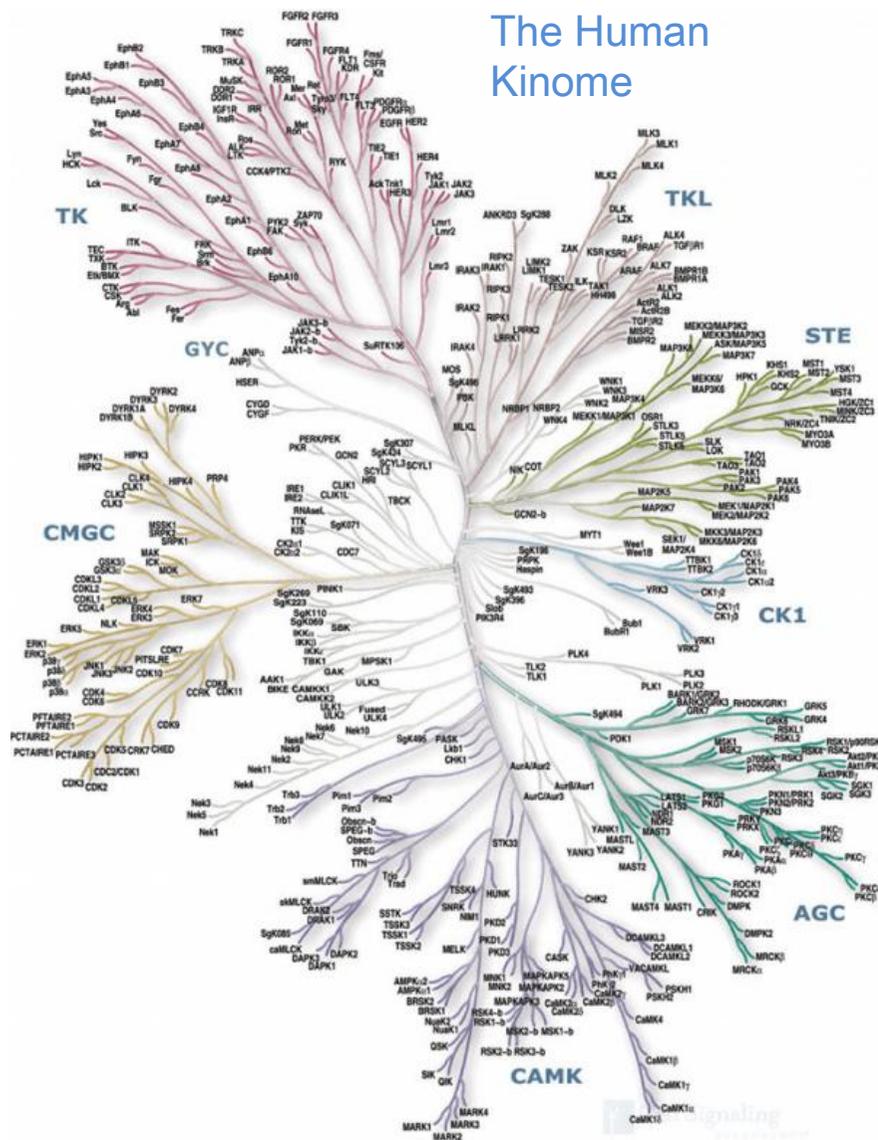
kinomes from other organisms and phylogenetic trees were used to group the protein kinases into the equivalent families. This analysis, based on sequence alignment, showed a number of differences compared with the human kinome (Previous Page - Figure 1-7). The protein kinase constitute up to 2% of the *T. brucei* genome, a similar proportion to the human kinome. In relative terms some families are expanded or reduced and some unique features also exist.

Of note is the complete absence of ePKs with sequences corresponding to tyrosine kinase or tyrosine kinase-like groups. This is despite the fact that tyrosine phosphorylation has been well documented in trypanosomes - with localisation and phosphoproteomic studies clearly demonstrating this (Parsons et al., 1991; Nett et al., 2009a). Several candidates for conducting tyrosine phosphorylation exist in the *T. brucei* kinome; these include several DYRK, CLK and STE7 protein kinases (which can act as atypical tyrosine kinases) as well as Wee1, which typically phosphorylates a conserved tyrosine in the ATP binding pocket of CDKs.

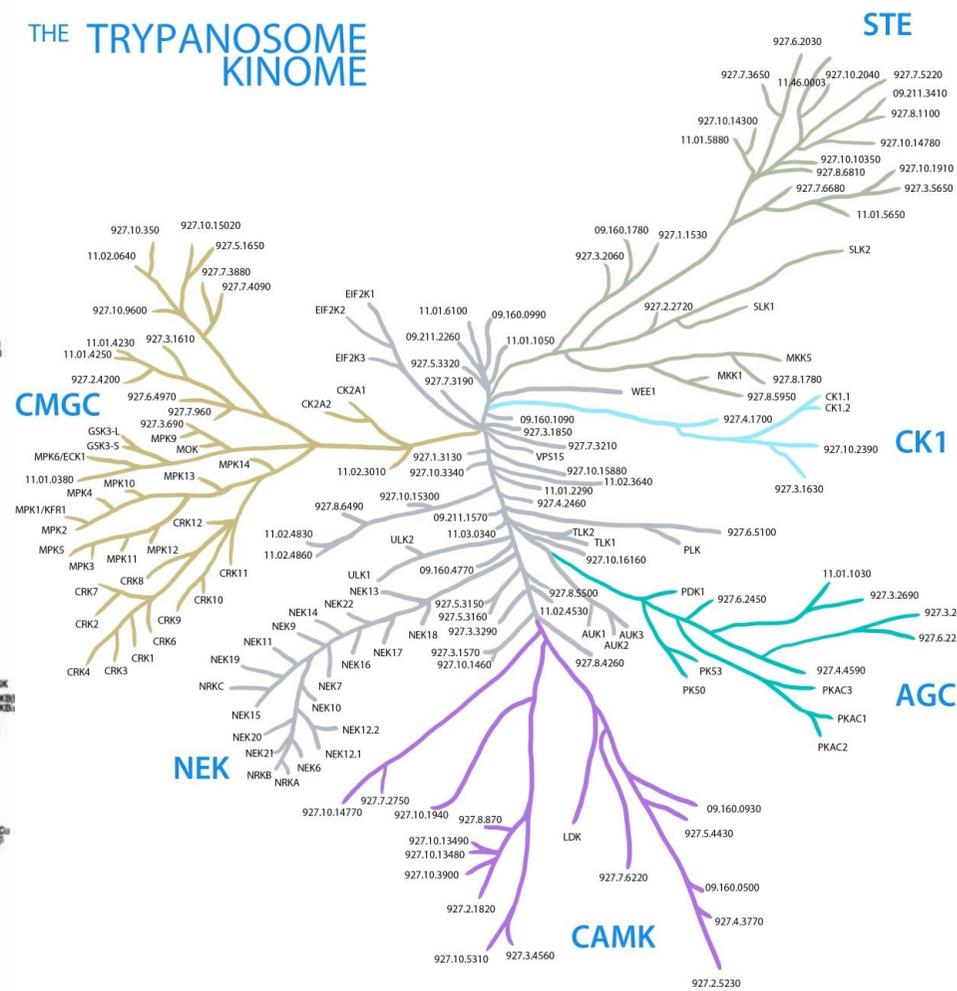
The AGC and CAMK families are two groups that are underrepresented in the *T. brucei* kinome. The CAMK family (in other organisms includes calcium/calmodulin regulated kinases and AMP-regulated Kinase) is small in *T. brucei* with only 13 members. The AGC kinase family is very small in *T. brucei* consisting of only 12 genes which makes it, proportionally, 50% smaller than the family in humans. Of the major families, the overrepresented groups are the CMGC kinases (42 genes) and STE kinases (25 genes). The CMGCs contain CRKs, 11 in total, 10 MAPKs with the remainder falling into smaller families such as the DYRKs, CLKs, GSKs, MAPK-like and SRPKs.

The STE kinases (named after sterile yeast mutants) contain the upstream activators of MAPKs. Two STE7 type kinases (which can function as MAPKKs) were identified with another 5 being very similar to these, suggesting they may function as STE7 kinases. Fourteen STE11 kinases (which can act as MAPKKKs) were identified but only 2 STE20 kinases (which in other organisms act as MAPKKKKs) were identified. The remainder of the STEs were uncategorised.

The Human Kinome



THE TRYPANOSOME KINOME



signaling
www.signaling.org

Previous Page - Figure 1-7: The Trypanosome Kinome – a stylised representation of the eukaryotic protein kinase complement of *Homo sapiens* and *T. brucei* based on the depiction of the human kinome by Manning *et al.* (2002). Protein kinase families are grouped together into families defined by amino acid sequence characteristics. In this illustration the node arrangements or length of the branches do not have any phylogenetic significance, and are purely illustrative. Number of ePKs depicted: *H. sapiens* N=518, *T. brucei* = 156.

The other group of expanded kinases, that actually number more in *T. brucei* than in humans, are the NEK kinases. *T. brucei* possesses 20 genes for NEK kinases compared to 15 in humans. These are discussed in more detail in Chapter 4. The study by Parsons *et al.* (2005) also identified 19 protein kinases that were very divergent from known families and thus could not be placed into the currently known categories, these have been termed Orphan or unique protein kinases.

During the analysis of the *T. brucei* phosphoproteome, researchers at the University of Dundee used different bioinformatic techniques to define the kinome of *T. brucei* and this led to the expansion of some of the groups by placing some of the orphan kinases into groups such as the AGC and STE families (Nett *et al.*, 2009b). This analysis also reduced the number of aPKs to 12.

Many studies have investigated the function of protein kinases in *T. brucei* using reverse genetics. However, none have so far identified the direct phosphorylation of a substrate *in vivo*, this is particularly troublesome in trypanosomes as very few reagents exist with which to probe the results of any perturbations in signalling pathways. For example, no antibodies exist for probing the phosphorylation state of any phosphosite in any *T. brucei* protein kinase and only a few exist for any phosphorylated proteins (phospho-PIP39, phospho-GPEET and phospho-PSSA2) (Bütikofer *et al.*, 1999; Fragoso *et al.*, 2009; Szöör *et al.*, 2010).

Other methods of investigating phosphorylation has included phosphoproteomics which has shown the relevance of phosphorylation in the both the bloodstream and procyclic form of the parasite (Nett *et al.*, 2009a, 2009b). MS analysis of phosphotyrosine-containing peptides isolated by immunoprecipitation of proteins from PCF whole cell lysates using anti-phosphotyrosine antibody detected 34 phosphotyrosine containing proteins. This study also localised phosphotyrosine using IFA and found the predominant staining to be associated with the flagellar

axoneme, basal body and cytoskeleton - this pattern was found in PCF and BSF parasites. Of the 34 proteins identified, 19 were protein kinases. Of these all but one were CMGC kinases (the other being a NEK), and included mainly MAPK or MAPK like protein kinases. The peptides of 8 MAPKs were fully phosphorylated in the TxY motif suggesting the same mode of activation as other canonical MAPK systems. The authors also suggested roles for tyrosine phosphorylation in regulating the CRKs CRK1, CRK2 and CRK3. A more comprehensive analysis of the bloodstream form cytosolic phosphoproteome fraction 1204 phosphopeptides across 491 proteins, of these 43 were ePKs (Nett et al., 2009b). The majority of these were CMGCs with AGCs, CAMKs, STEs and NEKs composing most of the remainder. Several "Other" kinases were detected, as well as the aPK TOR4. Serine was the most common residue phosphorylated followed by threonine and then tyrosine (63%, 24% and 13% respectively). For non-protein kinase phosphosites the ratio was much different at 75%, 21.5% and 3.5% respectively. This indicates the importance of tyrosine phosphorylation in regulating protein kinases - although most of those detected to contain tyrosine phosphorylation were CMGC kinases. The authors compared phosphosite prediction algorithms for other eukaryotes against the data set and found 25% of phosphosites were not predicted by these algorithms, which appear to indicate that *T. brucei* possesses protein kinases with novel substrate specificity. The results of kinome predictions and phosphoproteomics suggest trypanosomes have, to some extent, canonical MAPK pathways. However, there are no G-proteins or receptor tyrosine kinases in *T. brucei*, suggesting that the activation of these pathways occurs in a non-canonical manner.

Many of these protein kinases have been investigated at a molecular level in trypanosomes but only a few have been thoroughly investigated as drug targets. Several of these are conserved protein kinases, found in both the parasite and human host but nevertheless have been genetically and chemically validated as drug targets. Aurora kinases are conserved regulators of the cell cycle and several studies have investigated the biological role of the Aurora kinase AUK1 of *T. brucei*. It has been identified to be inhibited by both experimental compounds (hesperadin, VX-680) and danusertib, a compound in phase II clinical trials (Li and Wang, 2006; Jetton et al., 2009). Derivatives of danusertib have been shown to exhibit selective toxicity to *T. brucei* BSF over

mammalian cells and this drug is now a proposed lead compound for developing trypanosome specific Aurora kinase inhibitors (Ochiana et al., 2012).

Glycogen synthase kinase 3 is another widely conserved protein kinase, with *T. brucei* possessing a long and a short homologue. When the short isoform is ablated by RNAi in BSF parasites a severe growth defect occurs. Recombinant GSK-3 short can be inhibited by inhibitors designed to target human GSK3- β and their *in vitro* activity is correlated to their potency against *T. brucei in vivo* (Ojo et al., 2008). Bioinformatic analysis and molecular modelling have suggested that there are exploitable structural differences between the parasite and human kinases (Osolodkin et al., 2011). The crystal structure of *L. major* GSK-3 short has also been determined and allowed modelling and SAR studies of GSK-3 short of *L. major* and *T. brucei* (Ojo et al., 2010). *TbGSK-3* has been used in a high throughput screen of a 16,000 compound library to detect inhibitory compounds. The hits from this screen were counterscreened against a panel of human protein kinases. They were further tested for efficacy against *T. brucei* BSF parasites and toxicity against MRC-5 cells, identifying several attractive lead compounds in the process. However, many of the compounds were not selective for the *T. brucei* GSK3-short over the human homologue (in the enzyme panel screen) indicating that selecting conserved protein kinases as drug targets may prove problematic when it comes to developing clinically useful inhibitors. However, taken together these studies make GSK3-short the most fully validated *T. brucei* protein kinase that has currently been studied.

The CRK3 protein kinase of *Leishmania* spp. has been investigated as a drug target. It was shown to be essential for proper cell cycle progression in gene knockout studies using *L. mexicana*. LmxCRK3 was epitope tagged, purified from *L. mexicana* promastigotes and used in *in vitro* kinases assays, where it was shown to be inhibited by flavopiridol, a CDK inhibitor. This compound also caused cell cycle arrest in *L. mexicana* promastigotes (Hassan et al., 2001). LmxCRK3 inhibitors from a 634 compound library were identified in screens performed using protein purified from promastigotes. Sixteen of these, from various chemical categories, were then shown to have anti-leishmanial activity against *L. donovani* amastigotes (Grant et al., 2004). Bacterial expression of CRK3 and CYCA (the cyclin partner of CRK3) allowed the reconstitution of protein kinase

activity *in vitro* from bacterial expressed protein, allowing functional biochemical studies to be performed (Gomes et al., 2010). This development allowed active CRK3:CYCA to be co-expressed and co-purified in large quantities. The development of a high-throughput screening assay format allowed large chemical libraries to be screened against the *L. major* CRK3:CYC6 complex. The number of compounds in the various libraries was in the region of 30,000 (Cleghorn et al., 2011; Walker et al., 2011b). The hits generated by these screens were investigated against promastigote and amastigote form *L. major* with some of the compounds showing low micromolar IC50 values. However, some of the compounds were also toxic towards the *Leishmania* infected macrophages used in these experiments, demonstrating undesirable off target effects. The CRK3 protein kinase of *Leishmania* is a genetically and chemically validated drug target highlighting the potential of investigating parasite protein kinases in this way.

1.6 Validating Drug Targets

The differences between the human and trypanosome kinomes and individual protein kinase may make parasite ePKs valuable new drug targets. Protein kinases have been demonstrated as druggable targets, with selective inhibition of targets being achieved within the human kinome. They are typically formulated as orally available treatments, and several, such as gefitinib and sorafenib, can cross the blood-brain barrier - these inhibitors can be used to treat brain tumours such as glioblastoma (Korfel and Thiel, 2007). This makes *T. brucei* protein kinase potential drug targets, where inhibitors could eventually be developed that satisfy the target product profile defined by the DNDi for new antitrypanosomal drugs (Table 1-1). In order to pursue a trypanosome protein kinase as a drug target, worthy expending cost and effort on, it must be validated as one. This involves showing a protein is essential to the parasite and can be inhibited using a drug like molecule (Wang, 1997). Similarly, potential drug targets can be assessed using a “traffic lights system” to indicate the readiness and suitability of a target to move along the drug development pipeline. Target product profiles define the desired characteristics of the intended product and can help with decision making processes throughout the drug development programme by providing criteria to exclude undesirable products (Frearson et al., 2007; Wyatt et al., 2011).

Essential protein kinases can be validated as essential by genetic and chemical methods to prime drug development pipelines with new targets. A target based approach typically builds on basic research that has sought to understand the biology of potential drug targets; the development of imatinib was a classic example of this. Target based approaches can absorb a lot of resources and funding, and very many of them fail, thus they can be risky strategies to undertake. However, they do generally build a body of knowledge on the target and the way the target is inhibited. Approaches to determine if a target is essential to an organism can consist of genetic validation or chemical validation, though a combination of these is usually required to provide confidence in the target for a full drug development programme. Genetic techniques usually used in *T. brucei* are gene knockout (by targeted allelic replacement) and RNA interference. Other techniques such as overexpression of a mutant protein to achieve dominant negative phenotypes can also be used. Complete gene knockout will show that a gene is not essential in the life cycle stage in which it is performed under culture conditions (Frearson et al., 2007). If a gene is essential a tetracycline inducible version of the gene can be incorporated prior to knockout of the endogenous alleles. Removal of tetracycline causes expression of the exogenous copy to be stopped, inducing the knockout phenotype (Roper et al., 2002). Viability *in vivo* should be examined using an animal model. RNAi is easy to perform but if no phenotype is observed it does not exclude the target from being essential, as RNAi can have varying penetrance depending on the system used and the stability of the protein encoded by the target mRNA (Kolev et al., 2011). Both techniques ultimately either remove or radically deplete a protein from the parasite and any phenotype observed may be due to disruption of protein complexes or other structural requirements. Thus other techniques are required to identify whether enzymatic activity is required for a given targets essentiality. This can be achieved in knockout parasites by re-expressing a non-active version of the protein. Expression of re-coded RNAi refractory mutants in RNAi cell lines is also feasible (Szöör et al., 2006).

Chemical validation of a target uses a drug or experimental inhibitor to show that specific inhibition of a target is lethal or detrimental to a pathogen (Cong et al., 2012). If this can be carried out in the appropriate life cycle stage or in a

model infection it increases the strength of the validation. It further strengthens the case for a target being druggable, by showing inhibitory compounds can enter the cell and can be used in counter screens against host cells. However, this strategy is not without its difficulties and very often a specific inhibitor will not exist for a target during the early stages of target validation/drug discovery. The problem of off-target effects can be difficult to establish and identifying compounds acting on a specific target can be challenging to achieve using a phenotypic screen. However, chemical-genetic approaches such as generating analogue sensitive kinase alleles (ASKA) can be used to generate mutant proteins that are highly sensitive to specific experimental compounds.

The drugs melarsoprol, suramin and pentamidine were developed using phenotypic screens in that they were detected to be effective at killing *T. brucei*. Eflornithine arose from studies to develop anti-cancer treatments and was eventually found to be effective against *T. brucei*. Eflornithine was rationally designed to inhibit ornithine decarboxylase, though the turnover of this enzyme in mammalian cells is so rapid that it is ineffective. The rate of turnover in *T. b. gambiense* is low enough that a selective toxicity is achieved. Eflornithine is the only anti-trypanosomal drug with a well understood mode of action and selectivity (Burri and Brun, 2003; Steverding, 2010). Phenotypic screens are making something of a resurgence due to, in some quarters, the perceived failure of the target based approach and the ability to use high-content screening to detect very detailed and specific phenotypic information. Other techniques such as chemical proteomics can be used to identify the target of compounds detected in phenotypic screens.

Several studies have attempted to chemically validate protein kinases as drug targets in *T. brucei*, in particular using chemical proteomics. These methods are becoming more informative due to improvements in mass spectrometry technology (Cong et al., 2012). Chemical proteomics has been used to identify the targets of several trypanocidal 2,4-diaminopyrimidines that were identified in high-throughput screens (HTS) against whole parasites (Mercer et al., 2011). This compound family was known to possess activity against human CDKs and MAPKs, but the hits identified in the HTS had much lower IC50s against *T. brucei* and *Leishmania* spp. than against L929 fibroblasts. One of the hits, SCYX-5070 was also used to cure experimental *T. brucei* infections in a mouse model, with

no apparent acute toxicity in this system. In order to determine the targets of the compound the inhibitor was coupled to resin beads and mixed with cell lysates; free inhibitor and ATP was then used to elute interacting proteins. Proteomic analysis of these proteins identified a number of MAPKs and CRKs. This analysis used MALDI-MS and the data generated was non-quantitative.

Advances in MS technology, combined with targeted enrichment strategies have allowed this type of analysis to go much further, and identify the potencies with which a given protein kinase inhibitor binds to an individual protein kinase. By pulling down protein kinases using beads coated with a broad spectrum of inhibitors (KinoBeads) in the presence of varying quantities of a single, free inhibitor the amount of each protein kinase can be determined using quantitative mass spectrometry (in this case ITRAQ). This allows the IC₅₀ of an inhibitor to be calculated for any protein kinases detectable in the assay. This, in theory, allows kinome-wide determination of inhibitor potency in the absence of enzyme panels that are often used to profile inhibitors to determine their target specificity. This was recently applied to the trypanosome kinome allowing the profiling of inhibitors of GSK3, PK50 and PK53. This also allowed the counterscreening of the human kinome from MRC5 cells to identify potential anti-targets (Urbaniak et al., 2012b). However, it does not detect any protein kinase (or other protein) that does not bind to the Kinobeads - in this case the inhibition of PK50/PK53 could not be determined as these and other human NDR kinases were not bound by the Kinobeads. This study showed that it was possible to achieve different inhibition profiles in the trypanosome kinome compared to the human kinome, suggesting that the differences between the host/pathogen kinomes could be enough to allow the development of a usable, therapeutic protein kinase inhibitor.

Table 1-1: Target-product profile for next generation antitrypanosomals as determined by DNDi (<http://www.dndi.org/diseases/hat/target-product-profile.html>) highlighting the ideal scenario, an acceptable scenario and the current first line therapy.

Ideal	Acceptable : improvement to current St2 Tx	NECT
Effective against stage 1 and 2	Effective against stage 1+2 (used stage 2 only)	stage 1+2 (used stage 2 only)
Broad Spectrum (<i>gambiense</i> and <i>rhodesiense</i>)	Efficacy against <i>gambiense</i> only	<i>gambiense</i>
Clinical efficacy > 95% at 18 months follow up	To be determined by expert consultation	clinical efficacy: 96.5% (ITT NECT Study)
Effective in melarsoprol refractory patients		effective
<0.1% drug related mortality	<1% drug related mortality	1.2% possibly related mortality (NECT Field)
Safe also during pregnancy, for breastfeeding women and children		no specific adverse event found in babies born or being breastfed after treatment (NECT Field)
Adult and paediatric formulations		DFMO paediatric dosing available + Nifurtimox 5 mg tablets to be cut
No monitoring for AEs	Weekly simple lab testing (field testing)	hospitalization required
< 7 days p.o. once daily (DOT)	10 days p.o. (up to tid)	7 days IV infusion (bid) + 10 days po (tid)
< 7 days i.m. once daily	10 days i.m. once daily	
Stability in Zone 4 for > 3 years	Stability in Zone 4 for > 12 months	Stability in Zone 4 for > 24 months
Cidal	Static	DFMO static + Nifurtimox cidal
Multitarget	Unique target (but not uptake via P2-transporter only)	
< 30 € / course* (only drug cost)	< 100 €* / course	222.5 € / course (in 4 treatments kits; WHO)
	< 200 €* / course ok if very good on other criteria	

* It is expected that donor agencies will pay this, not patients. Considering that some 20-50,000 patients per year might require treatment, this is still realistic.

1.7 Research Aims

The work presented in this thesis aimed to genetically validate the protein kinase complement of *T. brucei* as drug targets, generating a library of RNAi cell lines for use in further studies. Several protein kinases were investigated in greater detail due to their potential amenability for rational drug design or their specific phenotype when ablated by RNAi. Due to the range of subjects covered in the course of these studies each results chapter contains a specific introduction and discussion in order to put the results in proper context. The generation of a kinome-wide library of RNAi plasmids and cell lines is presented first, second the examples of protein kinases interesting as drug targets and third the observations on a protein kinase involved in regulating bloodstream to procyclic form differentiation.

2 Materials and Methods

2.1 Bioinformatics

2.1.1 Sequence Retrieval

DNA and amino acid sequences for genes of interest were retrieved from TriTrypDB in FASTA format (<http://tritrypdb.org/tritrypdb/>). In certain instances annotated amino acid sequences were obtained from Genbank in GenPept format (<http://www.ncbi.nlm.nih.gov/genbank/>). TriTrypDB was also used to view collated data available for genes of interest including protein domain architecture, transcriptomic and proteomic evidence and other post-translational modifications.

2.1.2 Sequence Manipulation

In silico manipulation of DNA and amino acid sequences was conducted using CLC Genomics Workbench (CLC Bio). Features of this software allowed the design of oligonucleotide primers, sequence alignment, in silico cloning, restriction digest mapping and reverse translation/recoding of proteins. This software allows PFAM searches to identify protein domains to be carried out within it (<http://pfam.sanger.ac.uk/>). Sequences were analysed using the full PFAM library downloaded as an add-on from CLC Genomics plugins (<http://www.clcbio.com/products/clc-genomics-workbench/>).

2.1.3 Molecular Modelling

Homology models of the NEK 12 kinases were constructed using Molecular Operating Environment. *Molecular Operating Environment (MOE)*, 2011.10; Chemical Computing Group Inc., 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2011. Access to this software and training in its use was kindly provided by Dr Paul Selzer and colleagues (Intervet) at the COST Action CM0081 Training School - In Silico Tools for Drug Target Prioritisation (University of Siena, 2011).

For each NEK12 isoform (Tb927.8.7110, Tb927.4.5310) the amino acid sequence was used to search the MOE inbuilt PDB database for templates then aligned

using MOE's search and alignment function. Alignment was manually checked to ensure key residues were correctly aligned. Modelling was carried out using 10 stochastic iterations using the forcefield AMBER99 R-Field (used for amino acids and to include the effect of solvation in water on the model) with the lowest energy model being selected. MOE models the backbone, followed by loops and sidechains, and chooses the optimum model. Each optimum model was validated by viewing the Ramachandran plot for all residues and any amino acid lying outside the optimal areas was selected and energy minimised to make the model more energetically favourable before exporting it as a .PDB file for further use.

Table 2-1: Template PDB structures used for homology modelling of *T. brucei* NEK 12.1 and NEK12.2

Template PDB ID	Description	Pubmed ID
2JAV	HsNEK2 PI Ligand	17197699
2W5H	HSsNEK2 Apo	19124027
2WQM	HsNEK7 Apo	19941817

2.1.4 Data Mining for Potential NEK 12 Inhibitors

To identify ligands with the potential to bind into the NEK 12 models, the two models were screened using Ligand Discover at Edinburgh University (LIDAEUS) Models containing a pyrrole-indolinone ligand were used to define the binding site that was then screened against the CAMLESICK 2 library, this contains 4.3 million commercially available compounds that fit the Lipinski Rules, and 3.2 of these meet Oprea guidelines for lead-likeness (Hann and Oprea, 2004; Taylor et al., 2008). Sitepoint matching tolerance was set to 0.04 Å (from 0.02 Å), which provides lower stringency but increases the number of returned compounds. The top 100, 000 compounds were returned as a .SDF file.

2.1.5 Virtual Screening

Compounds returned from LIDAEUS screening were docked into the NEK 12 models using Autodock (rigid docking) and Vina (flexible sidechain docking) using a PyMOL plugin under the supervision of Dr D. Houston, University of Edinburgh (Seeliger and De Groot, 2010). Proteins and ligands were visualized using the PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC. The ligand

from the model template 2JAV was again used to define the centre of the active site box for virtual screening (the volume of the molecule where the software will attempt to dock the compounds). The top 500 compounds were ranked based on their binding energy and the top 100 manually validated. Compounds ranking in the top 100 that exhibited a posematch (when both docking techniques predict the same binding conformation) were classed as being interesting targets for inhibitor studies and the top 10 for each were investigated for purchase.

2.2 Molecular Biology

2.2.1 DNA Sequencing

DNA sequencing was performed by DNA Sequencing & Services (MRCPPU, College of Life Sciences, University of Dundee, Scotland, www.dnaseq.co.uk) using Applied Biosystems Big-Dye Ver 3.1 chemistry on an Applied Biosystems model 3730 automated capillary DNA sequencer.

2.2.2 Polymerase Chain Reaction

Oligonucleotide primers were designed using CLC Genomics Workbench and synthesised by Eurofins MWG Operon (Ebersber, Germany).

Standard PCRs (e.g. colony PCR): template DNA (100 ng genomic DNA, 100pg plasmid DNA or bacterial inoculation), 2.5 µl of 10 x PCR mix (1.13 mg ml⁻¹ BSA, 450 mM Tris pH 8.8, 110 mM ammonium sulphate, 45 mM MgCl₂, 68.3 mM β-mercaptoethanol, 44 µM EDTA pH 8.0, 10 mM dCTP, 10 mM dATP, 10 mM dGTP, 10 mM dTTP), 5 pmol of each primer, 1 unit of *Taq* polymerase and sterile distilled deionised water to a final volume of 25 µl. Thermocycling was programmed for an initial denaturation step of 96 °C for 5 minutes, followed by 30 cycles of: denaturation 96 °C for 30 seconds; annealing at [primer specific T_m] for 30 seconds; extension at 72 °C for *Taq* polymerase, 30 seconds per 0.5 kb of sequence. PCR conditions were subject to optimisation for specific primer pairs/templates and primer annealing temperature and elongation time were adjusted as required.

For generating PCR products for RNAi inserts or molecular cloning a proofreading enzyme was used, in this study Phusion (New England Biolabs) was used exclusively due to its robustness and reliability. Phusion reactions were set up as follows; Phusion HF buffer at 1x (stock at 5x), each dNTP at 200 μM , primers at 0.5 μM each, template DNA and Phusion polymerase at 0.02 U/ μl . Reactions were usually set up to 50 μl .

2.2.3 Restriction Enzyme Digests

All restriction endonucleases used in this study were sourced from New England Biolabs (NEB) and used according to the manufacturer's instructions and buffers. For restriction mapping of plasmids or generating fragments/backbones for cloning reactions were allowed to proceed for 1 hour before agarose gel electrophoresis. Plasmids that were being linearized in 5 μg amounts prior to transfection were generally digested overnight to ensure complete digestion.

2.2.4 Agarose Gel Electrophoresis

DNA sequences in this study were analysed by agarose gel electrophoresis which also serves as a method for purifying DNA fragments by size. UltraPure agarose (Invitrogen) was dissolved in 0.5 x TBE buffer (20mM Tris, 20 mM boric acid, 0.5mM EDTA, pH 7.2) at 1% w/v. For resolving larger fragments this was reduced to 0.8% w/v. After allowing the gel to cool SYBR-safe DNA stain (Invitrogen) was added to allow the visualization of DNA. Gels were then cast in the required size of gel support. DNA was prepared by addition of 6x Loading Buffer (0.25 % (w/v) bromophenol blue, 0.25 % (w/v) xylene cyanol FF, 30 % (v/v) glycerol, in H_2O) and loaded into wells. Gels were run at 120 V for 60 min depending on expected fragment sizes. Gels were imaged using a GelDoc (BioRad) and the associated Quantity One Software (BioRad).

In order to purify DNA fragments from agarose gels the DNA was visualised using a DarkReader blue light transilluminator, the band excised using a sterile scalpel blade and extracted from the agarose gel with a Gel Extraction kit (Qiagen) for general purposes, or a MinElute Kit (Qiagen) for more sensitive applications such as Gateway cloning. This was performed on a Qiacube (Qiagen) according to the manufacturer's instructions.

2.2.5 Ligations

When cloning constructs by the cohesive end technique compatible DNA fragments were ligated together using T4 DNA ligase (NEB) and the 10x Ligase Buffer supplied with it. Insert and plasmid backbone DNA fragments were mixed in equimolar amounts in a 10 μ l volume of 1xLigase buffer plus T4 DNA Ligase. These were incubated overnight at 16 °C. Some inserts were first subcloned into p-GEMT Easy (Promega) or PCR-Script (Stratagene) depending on the size of the insert. PCR-Script was typically used for inserts over 2 Kb due to its greater efficiency for larger inserts.

2.2.6 Gateway Cloning of RNAi Constructs

To recombine RNAi target sequences into the RNAi vector (pGL2084), 50 ng of purified, attB-tagged PCR product was mixed with 150 ng pGL2084 and adjusted to 8.75 μ l using TE buffer (pH 8.0) then mixed with 0.25 μ l of BP Clonase II enzyme (Invitrogen). Reactions were incubated at room temperature for 40 - 60 min and terminated by the addition of 1 μ l proteinase K solution (Ambion). A 1 μ l aliquot of the BP reaction mix was added to 50 μ l of Max Efficiency DH5 α *E. coli* (Invitrogen) which were incubated on ice for 30 min followed by 42°C heat shock for 30 seconds and recovery in SOC media for 1 hour at 37°C. The full transformation was plated onto agar plates containing 100 μ g ml⁻¹ Ampicillin (Sigma) and incubated overnight at 37°C. A colony was picked for overnight culture and plasmid DNA was extracted by Qiaquick Miniprep (Qiagen). Completed vectors were termed PTL###. An aliquot of 5 μ g of plasmid DNA was digested overnight with Ascl, ethanol precipitated and used to transfect *T. b. brucei* 2T1 BSF cells. Details of the RNAi library oligonucleotide, plasmid and cell line identification numbers are contained in supplemental data file S1 on the appended CD-ROM.

2.2.7 Ligation Independent Cloning of Kinase CDS for Protein Expression

PCR fragments containing full CDS for kinase genes or fragments thereof were cloned into the commercial pET-30Xa/LIC or pET-32Xa/LIC (Novagen) vectors following the manufacturer's instructions. pET-30Xa/LIC encodes an N-terminal

6xHis tag which can be cleaved from the protein of interest using activated Factor X. pET-32Xa/LIC encodes a N-terminal thioredoxin tag prior to the 6xHis tag that is intended to increase the solubility of the endogenous protein it is fused to. Expression of cloned genes is driven by the strong T7 promoter and the plasmids are for use with bacteria containing the DE3 lysogen. Oligonucleotides for amplifying the CDS of interest contained 20 nt specific to the gene and 5' of this sequence contained a 15 nt (GGTATTGAGGGTCGC-GOI) or 17 nt (AGAGGAGAGTTAGAGCC-GOI) vector specific sequence on the forward and reverse primers respectively. The purified PCR products were then treated with T4 DNA polymerase to remove 12 or 15 nt in a 3' to 5' direction leaving single-stranded overhangs allowing it to be annealed to the pre-treated vector and remain in frame. This annealed plasmid was then transformed into Gigasingle Competent Cells where the DNA was covalently bound and replicated.

A number of protein kinase sequences were analysed at the University of Edinburgh to define regions around the protein kinase domain optimised for protein expression and solubility (J. Cowman). This information allowed them to be cloned as kinase domain fragments for expression testing.

Table 2-2: List of oligonucleotides used to generate RNAi constructs prior to the establishment of a Gateway adapted cloning system.

OL number	Gene ID	Gene	Sequence	Restriction Sites
OL 3328	Tb927.7.3580	NEK11	GATCCTCGAGGGTACCCCTGAAGCTGGGAGAGCTTTG	XhoI, KpnI
OL 3329	Tb927.7.3580	NEK11	GATCTCTAGACTTAAGTTTTGTTATTACCTCGGGCG	XbaI, AflII
OL 3332	Tb10.70.2210	CRK3	GATCGGGCCCCGTACCAAAGGCTCTCGAGAAGAGGG	ApaI, KpnI
OL 3333	Tb10.70.2210	CRK3	GATCAAGCTTCTTAAGAGCGCTGGTAACACCTGAGT	HindIII, AflII
OL 3463	Tb10.61.0100	Orphan	GATCCTCGAGCCCCGGGAAATGGACTTTACCTCCGC	XhoI, XmaI
OL 3464	Tb10.61.0100	Orphan	GATCTCTAGAGGATCCCTGCCTCGGCTAATTCTGAC	XbaI, BamHI
OL 3465	Tb927.3.3290	Orphan	GATCCTCGAGCCCCGGGTGTATTGCATGTGGGAGGAA	XhoI, XmaI
OL 3466	Tb927.3.3290	Orphan	GATCTCTAGAGGATCCGGCACTTGTATACCACGCCT	XbaI, BamHI
OL 3467	Tb11.02.4530	Orphan	GATCCTCGAGCCCCGGGATCTCCAGACCCACTGATGG	XhoI, XmaI
OL 3468	Tb11.02.4530	Orphan	GATCTCTAGAGGATCCAGGCAAGGGCACATACAATC	XbaI, BamHI
OL 3598	Tb927.8.7110	NEK12.1	GGGGACAAGTTTGTACAAAAAGCAGGCTCTCGAGGGTACCTTGGGTCAAGGAAGTTTTGGAAGCG	XhoI, KpnI sites (+AttB1)
OL 3599	Tb927.8.7110	NEK12.1	GGGGACCACTTTGTACAAGAAAGCTGGGTTCTAGAGGATCCACCAAGAGCCACATTTCACTCTTT	XbaI, BamHI sites (+AttB2)
OL 3606	Tb927.3.3080	NEK6	GATCCTCGAGGTACCGAAGAAGTTGCAGGTCAGGC	XhoI, KpnI sites
OL 3607	Tb927.3.3080	NEK6	GATCTCTAGAGGATCCTCATCAAGGAAACTCGACCC	XbaI, BamHI sites
OL 3608	Tb927.3.3190	NEK7	GATCCTCGAGGGTACCGCCCTCGGTGTCATTCTTTA	XhoI, KpnI sites
OL 3609	Tb927.3.3190	NEK7	GATCTCTAGAGGATCCTATGCTCTCCCGATTCATCC	XbaI, BamHI sites
OL 3610	Tb927.8.1670	NEK13	GATCCTCGAGGGTACCCACGTGAAGAGAAGCAGCAG	XhoI, KpnI sites
OL 3611	Tb927.8.1670	NEK13	GATCTCTAGAGGATCCAATTTGACTGAAACGTCCG	XbaI, BamHI sites
OL 3612	Tb09.160.3480	TbPDK1	ATCCTCGAGGGTACCGGTCCGCAGGACCTTATGTA	XhoI, KpnI sites
OL 3613	Tb09.160.3480	TbPDK1	GATCTCTAGAGGATCCACCACAGGCCAGTAATCAG	XbaI, BamHI sites
OL 3614	Tb927.3.2240	AGC	GATCGGGCCCCCGGGGCGAACTCATTGTTGGTT	ApaI, XmaI
OL 3615	Tb927.3.2240	AGC	GATCTCTAGAGGATCCTTGTTCCTTTGTGCAGC	XbaI, BamHI
OL 3730	Tb927.6.2250	TbRAC	GATCTCTAGAGGATCCCTCCAGAGTTTCTGCTTGGG	XbaI, BamHI
OL 3731	Tb927.6.2250	TbRAC	GATCCTCGAGCCCCGGGCCACTACTTCATCGTCCGGT	XhoI, XmaI

Table 2-3: List of RNAi plasmids generated prior to the generation of a Gateway adapted vector system.

Plasmid	Gene ID	Gene Name	Backbone	Description
PGL 1987	Tb10.70.2210	TbCRK3	pGL 1875	Stem loop RNAi Vector AmpR
PGL 1988	Tb11.01.8460	TbCYC6	pGL 1875	Stem loop RNAi Vector AmpR
PGL 1989	Tb927.7.3580	NEK 11	pGL 1875	Stem loop RNAi Vector AmpR
pGL 2028	Tb10.61.0100	Orphan 0100	pGL 1875	Stem loop RNAi Vector AmpR
pGL 2029	Tb927.3.3290	Orphan 3290	pGL 1875	Stem loop RNAi Vector AmpR
pGI 2030	Tb11.02.4530	Orphan 4530	pGL 1875	Stem loop RNAi Vector AmpR
pGL 2051	Tb927.3.3080	NEK 6	pGL 1875	Stem loop RNAi Vector AmpR
pGL 2052	Tb927.8.1670	NEK 13	pGL 1875	Stem loop RNAi Vector AmpR
pGL 2053	Tb09.160.3480	TbPDK1	pGL 1875	Stem loop RNAi Vector AmpR
pGL 2064	Tb927.8.7110	NEK 12	pGL 1875	Stem loop RNAi Vector AmpR
pGL2068	Tb927.3.2440	AGC	pGL 1875	Stem loop RNAi Vector AmpR
pGL 2093	Tb927.3.3190	NEK7	pGL 1875	Stem loop RNAi Vector AmpR
pGL2084	N/A	N/A	pGL1875	Gateway adapted stem-loop RNAi plasmid backbone AmpR

Table 2-4: List of oligonucleotide primers generated in this study for the purposes of ligation independent cloning into the pET-30/32 Xa/LIC systems.

Oligo	Gene ID	Gene	Sequence	Region of Gene
OL 3158	Tb927.5.3210	Orphan	GGTATTGAGGGTCGCATGGAATGCGCAATGTTTCTTCCCT	CDS
OL 3159	Tb927.5.3210	Orphan	AGAGGAGAGTTAGAGCCTCAAGTCATGCATCAACGGCAAC	CDS
OL 3160	Tb927.7.3580	NEK 11	GGTATTGAGGGTCGCATGGATAGGTTTACTAAGGTCCGT	CDS
OL 3161	Tb927.7.3580	NEK 11	AGAGGAGAGTTAGAGCCTCAAGTAGCATTCCGCCTT	CDS
OL 3269	Tb11.02.4530	Orphan	GGTATTGAGGGTCGCATGGAACGAGACGATGGTG	CDS
OL 3270	Tb11.02.4530	Orphan	AGAGGAGAGTTAGAGCCCTAAGCATTGCTCCTGATAGA	CDS
OL 3271	Tb927.3.1570	Orphan	GGTATTGAGGGTCGCATGAGTGGCCATTGGGAGGG	CDS
OL 3272	Tb927.3.1570	Orphan	AGAGGAGAGTTAGAGCCTCACTCGTACACAATCTTTA	CDS
OL 3273	Tb927.3.3290	Orphan	GGTATTGAGGGTCGCATGTTACGACAATGGAGAG	CDS
OL 3274	Tb927.3.3290	Orphan	AGAGGAGAGTTAGAGCCTCACTTCTTCTCAGAAGTA	CDS
OL 3282	Tb09.160.0450	NEK 15	GGTATTGAGGGTCGCATGGATGGCGTGTGCG	CDS
OL 3283	Tb09.160.0450	NEK 15	AGAGGAGAGTTAGAGCCTCACTTGACAATCTTCTGCGA	CDS
OL 3284	Tb11.01.6650	NEK 21	GGTATTGAGGGTCGCATGGGAAACGGTGTGTGA	CDS
OL 3285	Tb11.01.6650	NEK 21	AGAGGAGAGTTAGAGCCCTACCCCATGTCAAGAGTAAACA	CDS
OL 3286	Tb927.3.3190	NEK 7	GGTATTGAGGGTCGCATGGCGCAGGTTGCAGAT	CDS
OL 3287	Tb927.3.3190	NEK 7	AGAGGAGAGTTAGAGCCCTACATCTTCTCCCTCGTACA	CDS
OL 3288	Tb11.01.2900	NEK 20	GGTATTGAGGGTCGCATGCACCGGGGGAGC	CDS
OL 3289	Tb11.01.2900	NEK 20	AGAGGAGAGTTAGAGCCTCACTCTAGGGACGAAATCGG	CDS
OL 3290	Tb10.70.0960	NEK 17	GGTATTGAGGGTCGCATGTCGTCAACGGATGCAA	CDS
OL 3291	Tb10.70.0960	NEK 17	AGAGGAGAGTTAGAGCCTCACTCAGCATCACCAATGC	CDS
OL 3292	Tb927.2.2120	NEK 22	GGTATTGAGGGTCGCATGGCTGATCTCCTTGAGCTTC	CDS
OL 3293	Tb927.2.2120	NEK 22	AGAGGAGAGTTAGAGCCTTAATTACTGCTGCCGCTGCT	CDS
OL 3294	Tb927.8.1670	NEK 13	GGTATTGAGGGTCGCATGGGGATGGATCACTACACCG	CDS
OL 3295	Tb927.8.1670	NEK 13	AGAGGAGAGTTAGAGCCTCACTGCACAGCAGGTGCGA	CDS

OL 3296	Tb927.8.7110	NEK 12.1	GGTATTGAGGGTCGCATGACCGCGCACAAACG	CDS
OL 3297	Tb927.8.7110	NEK 12.1	AGAGGAGAGTTAGAGCCTCAATGCGATATGGCTTTCTG	CDS
OL 3393	Tb927.3.3190	NEK 6	GGTATTGAGGGTCGCGGATCCATGAAGCCGCGGGATATT	CDS
OL 3394	Tb927.3.3190	NEK 6	AGAGGAGAGTTAGAGCCGTCGACTCAACACAACCCATAATTTGTAACC	CDS
OL 3395	Tb10.61.0100	Orphan	GGTATTGAGGGTCGCGGATCCGCGTGCAGACTTGCAATGAGCGACA	CDS
OL 3396	Tb10.61.0100	Orphan	AGAGGAGAGTTAGAGCCGTCGACTTAACTTCTCAGACTACAATGC	CDS
OL3552	Tb09.160.3480	TbPDK1	GGTATTGAGGGTCGCGAATTCATGTCAGCGTCCCGGATGGTGCCGG	CDS
OL3553	Tb09.160.3480	TbPDK1	AGAGGAGAGTTAGAGCCGTCGACTCAATCCCCCTAATCTTGCCAAG	CDS
OL 3695	Tb927.3.2440	AGC	GGTATTGAGGGTCGCATGCTCAATAGAGACTA	CDS
OL 3696	Tb927.3.2440	AGC	AGAGGAGAGTTAGAGCCCTAACTACCACCACC	CDS
OL 3697	Tb927.6.2250	TbRAC	GGTATTGAGGGTCGCATGACCATCGATTAC	CDS
OL 3698	Tb927.6.2250	TbRAC	AGAGGAGAGTTAGAGCCCCTATGTAGGTCTTGA	CDS
OL 3746	Tb927.4.5310	NEK12.2	GGTATTGAGGGTCGCATGACGGAACATAAGCGTCC	CDS
OL3747	Tb927.4.5310	NEK12.2	AGAGGAGAGTTAGAGCTCAACGAGGCACGGCATTCT	CDS
OL3916	Tb11.01.5650	DFK	GGTATTGAGGGTCGCTCTAGTGATTGTAACCTGA	Kinase Domain
OL3917	Tb11.01.5650	DFK	AGAGGAGAGTTAGAGCCTCACAAACAAAATGCATATG	Kinase Domain
OL3918	Tb11.01.5650	DFK	GGTATTGAGGGTCGCCTTTTTCTAACCGATTCCAA	Kinase Domain
OL3919	Tb09.211.2260	ZFO	GGTATTGAGGGTCGCTTATCCCTGGAGGCGTGCGA	Kinase Domain
OL3920	Tb09.211.2260	ZFO	AGAGGAGAGTTAGAGCCTTAGCGCCTATCGCCCCATTCCG	Kinase Domain
OL3921	Tb11.02.2050	ULK	GGTATTGAGGGTCGCCCCAAAATTCGGAATATGGA	Kinase Domain
OL3922	Tb11.02.2050	ULK	AGAGGAGAGTTAGAGCCTTATACCGCCATCAACCACTTCT	Kinase Domain
OL3923	Tb927.4.3420	Wee1	GGTATTGAGGGTCGCCGTCGCATTCTGACTGACTA	Kinase Domain
OL3924	Tb927.4.3420	Wee1	AGAGGAGAGTTAGAGCCTTATAACAACCCACTGCGACATAA	Kinase Domain
OL3925	Tb11.46.0003	STE	GGTATTGAGGGTCGCGCGTGCAACTTTAAGTTCAA	Kinase Domain
OL3926	Tb11.46.0003	STE	AGAGGAGAGTTAGAGCCTTAGGAATAGCTTTGCGAAAAGA	Kinase Domain
OL3927	Tb927.10.2040	STE	GGTATTGAGGGTCGAGATATTCCACATATCAGTT	Kinase Domain
OL3928	Tb927.10.2040	STE	AGAGGAGAGTTAGAGCCTTATAACAATGTTCTCCGCAAGGA	Kinase Domain

OL3929	Tb927.10.5140	MAPK2	GGTATTGAGGGTCGCCACATATCCTCCGAAATA	Kinase Domain
OL3930	Tb927.10.5140	MAPK2	AGAGGAGAGTTAGAGCCTTAGTGAAAGCCGCCACATACG	Kinase Domain
OL3931	Tb09.160.0570	Aur	GGTATTGAGGGTCGCATGTGGTCTTTGGACGATTT	Kinase Domain
OL3932	Tb09.160.0570	Aur	AGAGGAGAGTTAGAGCCTTAGTAACGGTTGCTTATGCGAC	Kinase Domain
OL3933	Tb927.6.4970	SRPK	GGTATTGAGGGTCGCGGCTACCACCCAGTCGTGGT	Kinase Domain
OL3934	Tb927.6.4970	SRPK	AGAGGAGAGTTAGAGCCTTAAAAAAGTGCTGAAAATTAT	Kinase Domain
OL3935	Tb11.02.3010	CMGC	GGTATTGAGGGTCGCGAGTTCAATTATGAGATTAG	Kinase Domain
OL3936	Tb11.02.3010	CMGC	AGAGGAGAGTTAGAGCCTAACATCCCTGATGAATGGAAC	Kinase Domain
OL3937	Tb927.8.7110	NEK12.1	GGTATTGAGGGTCGCTGCAGGTATGTCAAAAAAAG	Kinase Domain
OL3938	Tb927.8.7110	NEK12.1	AGAGGAGAGTTAGAGCCTTAGTCTAAGCCACGGCGAATGA	Kinase Domain
OL3939	Tb927.4.5310	NEk12.2	GGTATTGAGGGTCGCTGCAAATATGTTAAGAAGAA	Kinase Domain
OL3940	Tb927.4.5310	NEk12.2	AGAGGAGAGTTAGAGCCTTATTCTAAGCCACGGCGAATGA	Kinase Domain
OL3941	Tb09.160.3480	TbPDK1	GGTATTGAGGGTCGCGTGCCGGCCGATTTTGAGCT	Kinase Domain
OL3942	Tb09.160.3480	TbPDK1	AGAGGAGAGTTAGAGCCTTATGCGAGACCATCAAAATCAA	Kinase Domain

Table 2-5: List of oligonucleotides generated for site directed mutagenesis of catalytic lysine residues in various protein kinases.

Oligo	Gene ID	Gene	Sequence	Mutation
OL3954	Tb927.8.7110	NEK 12.1	CGATCTTTCAGCAATGGTGATGGATACAAACAATATGAGC	K70M
OL3955	Tb927.8.7110	NEK 12.1	GTCATATTGTTTGTATCCATCACCATTGCTGCAAAGATCG	K70M
OL3956	Tb9274.5310	NEK 12.2	GCTAATTTTTGCAGCAATGGTGATGGACACAAACAATATGAGC	K70M
OL3957	Tb9274.5310	NEK 12.2	GTCATATTGTTTGTGTCCATCACCATTGCTGCAAAAATTAGC	K70M
OL4057	Tb11.01.5650	DFK	CTCAGGTCGCCATCATGTGCATTCCGCGGTTG	K950M
OL4058	Tb11.01.5650	DFK	CAACCGCGGAATGCACATGATGGCGACCTGAG	K950M
OL4059	pGL2218	NEK12.1recoded	GTATTTTTGCCGCCATGGTAATGGACACTAAC	K70M
OL4060	pGL2218	NEK12.1recoded	GTTAGTGTCATTACCATGGCGGCAAAAATAC	K70M
OL4061	pGL2222	NEK12.2recoded	GGACTTATATTTGCAGCGATGGTAATGGATACGAACAAC	K70M
OL4062	pGL2222	NEK12.2recoded	GTTGTTCGTATCCATTACCATCGCTGCAAATATAAGTCC	K70M

Table 2-6: List of plasmids generated for the purpose of protein expression in *E. coli* DE3 lysogen cell lines.

Plasmid	Gene ID	Gene Name	Backbone	Description	Region of Gene
PGL 1990	Tb927.3.3080	NEK 6	pET-32 Xa/LIC	N terminal TRX, S, His tags. AmpR	CDS
PGL 1991	Tb927.3.3190	NEK 7	pET-32 Xa/LIC	N terminal TRX, S, His tags. AmpR	CDS
PGL 1992	Tb927.7.3580	NEK 11	pET-32 Xa/LIC	N terminal TRX, S, His tags. AmpR	CDS
PGL 1993	Tb927.8.7110	NEK 12.1	pET-32 Xa/LIC	N terminal TRX, S, His tags. AmpR	CDS
PGL 1994	Tb927.8.1670	NEK 13	pET-32 Xa/LIC	N terminal TRX, S, His tags. AmpR	CDS
PGL 1995	Tb09.160.0450	NEK 15	pET-32 Xa/LIC	N terminal TRX, S, His tags. AmpR	CDS
PGL 1996	Tb10.70.0960	NEK 17	pET-32 Xa/LIC	N terminal TRX, S, His tags. AmpR	CDS
PGL 1997	Tb11.01.2900	NEK 20	pET-32 Xa/LIC	N terminal TRX, S, His tags. AmpR	CDS
PGL 1998	Tb11.01.6650	NEK 21	pET-32 Xa/LIC	N terminal TRX, S, His tags. AmpR	CDS
PGL 1999	Tb927.2.2120	NEK 22	pET-32 Xa/LIC	N terminal TRX, S, His tags. AmpR	CDS

PGL 2000	Tb10.61.0100	Orphan 0100	pET-32 Xa/LIC	N terminal TRX, S, His tags. AmpR	CDS
PGL 2001	Tb927.3.3290	Orphan 3290	pET-32 Xa/LIC	N terminal TRX, S, His tags. AmpR	CDS
PGL 2002	Tb11.02.4530	Orphan 4530	pET-32 Xa/LIC	N terminal TRX, S, His tags. AmpR	CDS
pGL 2023	Tb09.160.3480	TbPDK1	pET-32 Xa/LIC	N terminal TRX, S, His tags. AmpR	CDS
pGL 2099	Tb927.8.7110	NEK12.2	pET-32 Xa/LIC	N terminal TRX, S, His tags. AmpR	CDS
pGL 2181	Tb09.211.2260	ZFO	pET-30 Xa/LIC	N terminal, S, His tags. KanR	Kinase Domain
pGL 2182	Tb11.02.2050	ULK	pET-30 Xa/LIC	N terminal, S, His tags. KanR	Kinase Domain
pGL 2183	Tb927.4.3420	Wee1	pET-30 Xa/LIC	N terminal, S, His tags. KanR	Kinase Domain
pGL 2184	Tb11.46.0003	STE	pET-30 Xa/LIC	N terminal, S, His tags. KanR	Kinase Domain
pGL 2185	Tb927.10.2040	STE	pET-30 Xa/LIC	N terminal, S, His tags. KanR	Kinase Domain
pGL 2186	Tb927.10.5140	MAPK	pET-30 Xa/LIC	N terminal, S, His tags. KanR	Kinase Domain
pGL 2187	Tb09.160.0570	Aur	pET-30 Xa/LIC	N terminal, S, His tags. KanR	Kinase Domain
pGL 2188	Tb927.6.4970	SRPK	pET-30 Xa/LIC	N terminal, S, His tags. KanR	Kinase Domain
pGL 2189	Tb11.02.3010	CMGC	pET-30 Xa/LIC	N terminal, S, His tags. KanR	Kinase Domain
pGL 2190	Tb927.8.7110	NEK12.1	pET-30 Xa/LIC	N terminal, S, His tags. KanR	Kinase Domain
pGL 2191	Tb927.4.5310	NEK12.2	pET-30 Xa/LIC	N terminal, S, His tags. KanR	Kinase Domain
pGL 2192	Tb09.160.3480	TbPDK1	pET-30 Xa/LIC	N terminal, S, His tags. KanR	Kinase Domain
pGL2197	Tb11.01.5650	STE -Differentiation Kinase	pET-30 Xa/LIC	N terminal, S, His tags. KanR	Kinase Domain
pGL2198	Tb11.01.5650	STE -Differentiation Kinase	pET-32 Xa/LIC	N terminal TRX, S, His tags. AmpR	Kinase Domain
pGL2199	Tb11.01.5650	STE -Differentiation Kinase	pET-32 Xa/LIC	N terminal TRX, S, His tags. AmpR	Kinase Domain
pGL2200	Tb927.8.7110	NEK12.1	pET-32 Xa/LIC	N terminal TRX, S, His tags. AmpR	Kinase Domain
pGL2201	Tb927.4.5310	NEK12.2	pET-32 Xa/LIC	N terminal TRX, S, His tags. AmpR	Kinase Domain
pGL2202	Tb09.211.2260	ZFO	pET-32 Xa/LIC	N terminal TRX, S, His tags. AmpR	Kinase Domain
pGL 2214	Tb927.8.7110	NEK 12.1	pET-32 Xa/LIC	N terminal TRX, S, His tags. AmpR	CDS K70M
pGL 2215	Tb927.4.5310	NEK12.2	pET-32 Xa/LIC	N terminal TRX, S, His tags. AmpR	CDS K70M

Table 2-7: List of oligonucleotides used to generated fragments for cloning of endogenous tagging vectors.

Oligo	Gene ID	Gene	Cloning into	Description	Sequence	Restriction Sites
OL 3774	Tb927.8.7110	NEK12.1	pGL1796	NEK12.1 Endogenous Tagging	CATGTCTAGAATGACCGCGCACAAACGTGC	XbaI
OL 3775	Tb927.8.7110	NEK12.1	pGL1796	NEK12.1 Endogenous Tagging	CATGGGATCCTCAATGCGATATGGCTTTCT	BamHI
OL3869	Tb09.211.2260	ZFO 5' UTR	pHG80	ZFO eGFP HA N-Ter Endogenous Tagging	GATCGGCGCGCCAATTTATCTGTGCAGCAGCG	Ascl
OL3870	Tb09.211.2260	ZFO 5' UTR	pHG80	ZFO eGFP HA N-Ter Endogenous Tagging	GATCGGATCCATCCAAAAAATGCAAGTGT	BamHI
OL3871	Tb09.211.2260	ZFO CDS	pHG80	ZFO eGFP HA N-Ter Endogenous Tagging	GATCTCTAGAATTGGCACAGTAGAACAAAT	XbaI
OL3872	Tb09.211.2260	ZFO CDS	pHG80	ZFO eGFP HA N-Ter Endogenous Tagging	GATCGGCGCGCCCCATCCAAATCGAAGACAAA	Ascl
OL3914	Tb11.01.5650	DFK 3' CDS	pGL2178	DFK 12Myc Endogenous tagging	GATCTCTAGACAACAAAAATGCATATGAAA	XbaI
OL3915	Tb11.01.5650	DFK 3' CDS	pGL2178	DFK 12Myc Endogenous tagging	GATCAAGCTTACGACGTCCGGATGGGGAGG	HinDIII

Table 2-8: List of plasmids generated for endogenous tagging of genes in this study.

Plasmid	Gene ID	Gene Name	Backbone	Description
pGL2111	Tb927.8.7110	NEK 12.1	pGL1796	pEnT6BG, AmpR. N-Ter GFP-TY-NEK 12.1. Linearise with AgeI.
pGL2167	Tb09.211.2260	ZFO	pHG 80	pEnT6BG derivative. AmpR. N-ter GFP 3HA ZFO UTR/CDS for endogenous taging Linearise with Ascl.
pGL2203	Tb11.01.5650	Diff STE	pGL2178	pNAT BSD AmpR 3' CDS of DFK for endogenous C-ter 12Myc tagging. Linearise with PstI.

2.2.8 Site Directed Mutagenesis

Mutations of single bases and deletions of multiple bases were attempted using the Quikchange Mutagenesis kit (Stratagene) according to manufacturer's instructions. In the Case of NEK12.1 and NEK12.2 the gene could not be mutated in this manner and a gene sewing approach was undertaken using a combination of the Quikchange mutagenesis primers and the pET-32 LIC primers. Mutations were carried out to change the catalytic lysine of kinase domains. This was determined in individual cases by sequence alignment prior to designing oligos to replace the endogenous lysine codon to a methionine codon.

2.2.9 Transformation of Bacteria

Chemically competent *E. coli* cells were transformed by heat shock. Plasmid DNA was added to 50 μ l competent cells and incubated on ice for 5 minutes before a 45 second heat shock at 42°C. Reactions were returned to ice for 2 minutes before cells were plated directly onto selective LB agar plates. Ligations and recombination reactions were incubated longer (60 min) prior to heat shock and reactions were recovered in 200 μ l of SOC medium for 60 min at 37°C with shaking prior to plating. When transforming with a plasmid carrying kanamycin resistance, cells were also recovered for 1 hour prior to plating.

2.2.10 Genomic DNA Extractions

Genomic DNA was prepared from cultured *T. brucei brucei* TREU 927 PCF forms using the DNeasy Blood and Tissue kit (Qiagen), following the Cultured Animal Cell protocol outlined in the manufacturer's instructions.

2.2.11 Preparation of DNA for Trypanosome Transfection

Five micrograms of a purified plasmid were digested with restriction endonucleases appropriate for linearising the plasmid, e.g. the pTL vectors were linearised with *Ascl*. Reactions were cleaned up by the addition of 0.1 volume of 3M sodium acetate pH 5.5 followed by 0.7 volumes of isopropanol then centrifugation at 13 000g x 20 min to precipitate and pellet DNA. DNA pellets were washed with 70% ethanol and dissolved in 10 μ l molecular biology grade H₂O under sterile conditions.

2.2.12 RNA Extraction, cDNA Preparation, and qPCR

For qPCR analysis of RNAi lines a total of 2×10^7 cells for induced and uninduced cultures were harvested at 24h post induction. Total RNA was isolated from the cell pellets using the Qiagen RNeasy Kit on the Qiacube (with on column DNaseI digest) to minimize RNase contamination. Samples were then treated again with RQ1 RNase-free DNaseI (Promega) to ensure gDNA was totally degraded. All pipetting equipment and gloves were cleaned with RNase-ZAP (Sigma Aldrich) prior to manipulating RNA samples. Sample concentration and purity were estimated using a NanoDrop UV Spectrometer.

To prepare cDNA, 2 μg of total RNA was used as a template in a reverse transcriptase reaction using the SuperScript Reverse Transcriptase III system (Invitrogen). RT-PCRs were primed with random hexamers (Invitrogen). RNaseOUT RNase inhibitor (Invitrogen) was included to reduce sample degradation. Following cDNA synthesis *E. coli* RNase H (Invitrogen) was added to degrade the template RNA. Samples were also processed that did not contain reverse-transcriptase to monitor for gDNA contamination in the qPCR template. qPCR reactions were set up using Applied Biosystems SYBR Green PCR master mix, each reaction was performed in quadruplicate and contained 12.5 μl Mastermix, 2.5 μl of each primer (3 μM), 5.5 μl H₂O and 2 μl of the template cDNA. Reactions were set up in MicroAmp® Optical 96-Well Reaction Plate (Applied Biosystems) and sealed with MicroAmp® Optical Adhesive Film (Applied Biosystems). Samples were analysed in an Applied Biosystems Prism 7500 using the default thermocycling settings with the reaction volume set to 25 μl . A denaturation step was added to check that only a single product was being formed by each primer pair.

Table 2-9: List of oligonucleotides used to prime qPCR reactions.

Oligo	Gene ID	Gene	Sequence
OL 3701	Tb09.160.3480	TbPDK1	GCTCCGCGAATCATATTACCA
OL 3702	Tb09.160.3480	TbPDK1	AGCAAGAATCACAAACCCTCTCA
OL 3703	Tb927.3.3080	NEK6	AAGCCGCGGGATATTTCC
OL 3704	Tb927.3.3080	NEK6	GTTACCTAAAAACAAGGCCTTCTCA
OL 3705	Tb927.8.1670	NEK13	AATTCATTGAGTCTTACCTGCAAAAG
OL 3706	Tb927.8.1670	NEK13	AGGTCCCTCCGCGATTAAAA
OL 3707	TbCRK3	CRK3	AAACAGCTCTTCGGGAGGTATCT
OL 3708	TbCRK3	CRK3	CCATCGGCACATATAACATCTAGAA
OL 3756	Tb927.8.7110	NEK12.1	TGGCGTTGAAACGGTGAAT
OL 3757	Tb927.8.7110	NEK12.1	GCGGTTGCCGCGTCTA
OL 3758	Tb927.8.7110	NEK12.1	GGAGTAGTGTAACGCCTCATCGT
OL 3759	Tb927.8.7110	NEK12.1	CTCGCAATCCTTCCAACCA
OL3896	Tb11.01.5650	DFK	CTCCATCTTCTCATCCTC
OL3897	Tb11.01.5650	DFK	CACCTCCATCTCCTTCTC
OL3898	Tb09.211.2260	ZFO	CCGAAGGAAAACGCTGAGTT
OL3899	Tb09.211.2260	ZFO	CCACTAGGCCACGACCATT
OL3900	Tb927.3.4860	MKK1	GGGAGGAAAATACTTGGGAGTGT
OL3901	Tb927.3.4860	MKK1	CGCCCCGGCCAATAAT
OL3902	Tb927.10.5270	MKK5	TCACGACGTGCATTCAAAGG
OL3903	Tb927.10.5270	MKK5	TGCCGCCGGAGTTCAT
OL3904	Tb927.4.3420	Wee1	TGACCGGCGCTGTAACCTTG
OL3905	Tb927.4.3420	Wee1	GTGCGCTGAACTGTGAAAGC
OL3906	Tb927.4.2500	PERK EIF	TGACGCGCCGCTCACT
OL3907	Tb927.4.2500	PERK EIF	GCCGACGCGGAATAAGC
OL3908	Tb11.02.2050	ULK1	TGTTTCCGGAGGGTAATGCT
OL3909	Tb11.02.2050	ULK1	CGGTGATGCATCTATTGTTTCTTG

2.3 Protein Production and Biochemistry

2.3.1 Bacterial Cell Lines

Single colonies of the respective *E. coli* strain were used to inoculate 2 ml of LB medium and incubated overnight at 37 °C. 0.5 ml of this culture was used to inoculate 100 ml LB broth which was cultured for several hours until an OD600 of 0.4 - 0.6 was reached. The culture was chilled on ice for 10 mins, then centrifuged at 1000 x g for 15 mins at 4 °C in 50 ml conical flasks (Falcon). The cell pellet was resuspended in 33 ml cold RF1 buffer (100 mM RbCl, 50 mM MnCl₂ · 4H₂O, 30 mM potassium acetate pH 7.5, 10 mM CaCl₂ (dihydrate), 15 % (w/v) glycerol, pH adjusted to 5.8 with 0.2 M acetic acid, filter-sterilised), incubated for 15 mins on ice and centrifuged again. The pellet was resuspended in 8 ml

cold RF2 buffer (10 mM MOPS pH 6.8, 10 mM RbCl, 75 mM CaCl₂·2H₂O, 15 % (w/v) glycerol, pH adjusted to 6.8 with NaOH, filter-sterilised) and incubated on ice for one hour. The suspension was divided into 50 µl single-use aliquots and flash-frozen in ethanol:dry-ice slurry or liquid nitrogen; aliquots were stored at -80 °C. Cell lines tested during the optimisation of protein expression included BL21 (DE3), BL21 (DE3) pLysS, and BL21 (DE3) Rosetta pLysS (Novagen). C41 (DE3) and C43 (DE3) were also tested (Dumon-Seignovert et al., 2004).

2.3.2 Protein Expression

The proteins examined in detail in this project were the NEK 12 kinases. The expression and purification for these was as follows. Rosetta pLysS cells were transformed with the appropriate plasmid and plated to yield single colonies. Single colonies were used to inoculate 50 ml starter cultures that were grown overnight until stationary. These were used to inoculate 1 litre main cultures (divided up into 200 ml aliquots in 1.5 litre conical flasks) that were grown at 37°C with shaking until an OD₆₀₀ of 0.6 was reached. Cultures were then cooled to 15 °C and then induced with 1 mM IPTG. Protein expression was allowed to continue overnight at 15 °C with shaking. Cells were harvested by centrifugation at 10 000 g for 15 min at 4 °C. Cell pellets were then lysed using BPER (Pierce) supplemented with Complete EDTA-Free protease inhibitors (Roche) and clarified by centrifugation at 50 000 g for 20 min at 4 °C.

Large-scale native purifications were performed using an IMAC (immobilised metal-ion affinity) column packed with MC-20 matrix (Poros) on an ÄKTA Purifier (GE Healthcare). The column was equilibrated with equilibration buffer (50 mM NaH₂PO₄, 300 mM NaCl, 0.5 mM imidazole, pH 8.0), the filtered soluble fraction of a bacterial extract was loaded onto the column, washed with 10 column volumes of wash buffer (50 mM NaH₂PO₄, 300 mM NaCl, 50 mM imidazole, pH 8.0) and eluted into several fractions with elution buffer (50 mM NaH₂PO₄, 300 mM NaCl, 500 mM imidazole, pH 8.0). Prior to kinase assays the eluted protein was buffer exchanged into 50 mM NaH₂PO₄, 300 mM NaCl, pH 8.0 using a PD-10 desalting column.

2.3.3 SDS-PAGE

Cell lysates and purified recombinant proteins were size-separated by SDS-PAGE (sodium dodecyl phosphate polyacrylamide gel electrophoresis). For general uses 12 % (w/v) polyacrylamide gels were used (or 8 % in the case of DFK analyses), with a layer of 5 % stacking gel to focus samples before separation in the resolving gel. Samples were prepared by addition of 2 x SDS-PAGE loading buffer (20 % (v/v) glycerol, 2.5 % (w/v) SDS, 0.05 % (w/v) bromophenol blue, 0.2 M Tris-HCl pH 6.8, 20 mM DTT) and denatured by boiling for 4 minutes. Electrophoresis was performed in XCell SureLock Mini Cell chambers (Invitrogen) with SDS-PAGE running buffer (2.5 mM Tris, 19.2 mM glycine and 0.01 % (w/v) SDS). The voltage was set at 100 V while samples were in the stacking gel, then increased to 180 V as samples progressed through the resolving gel. With each gel, 2 µg of protein marker (Pre-stained Protein Marker Broad Range, New England Biolabs or Precision Plus Dual Extra, BioRad) was run as a reference to estimated molecular weight of protein samples.

2.3.4 Coomassie Staining of SDS-PAGE Gels

SDS-PAGE gels were stained by immersion in Coomassie stain solution (2.5 % (w/v) Coomassie Brilliant Blue R-250, 45 % (v/v) methanol, 10 % acetic acid) for 1 hour at room temperature. Gels were destained with several changes of destain solution (10 % (v/v) methanol, 10% (v/v) acetic acid) overnight at room temperature with inclusion of tissue paper to adsorb free dye, following which they were equilibrated in distilled water and photographed on the Gel-Doc 2000 (Biorad) or dried for further analysis e.g. autoradiography.

2.3.5 Western Blotting

To analyse by Western blot, samples were separated by SDS-PAGE and transferred by 'wet' transfer to PVDF Biotrace membrane (Pall) in an X-Cell II Blot Module (Invitrogen). PVDF membrane was activated in methanol for 30 s, rinsed in ddH₂O then equilibrated in transfer buffer (20 % (v/v) methanol, 20 mM Tris, 15 mM glycine). The transfer chamber was assembled with all pads and filter papers soaked in transfer buffer, then filled with transfer buffer. Transfer was performed at 30 volts for 2 hours with the outer chamber filled with water to provide a coolant. After transfer the membrane was blocked with 5 %

powdered milk (Marvel) in TBST buffer (25 mM Tris-HCl pH 8, 120 mM NaCl, 0.1 % Tween-20) for at least 1 hour (often overnight at 4°C) followed by incubation with the appropriate antibodies diluted in blocking solution. Primary antibodies diluted in blocking solution were applied to the blot for 1 hour at room temperature then washed 3x 10 minutes in TBST. Secondary antibodies were applied in blocking buffer for 1 hour at room temperature. All secondary antibodies used were conjugated to horseradish peroxidase (HRP) and antibody labelling of protein was revealed by applying an ECL (enhanced chemiluminescence) substrate (ECL prime, Amersham; Illuminati Classico or Crescendo, Millipore) to the membrane according to manufacturer's instructions. Chemiluminescence was visualised by exposing X-Ray film to the blot and developing this in a Kodak Xomat film processor.

Blots were stripped with by incubation in stripping buffer (glycine 15 g/L, SDS 1g/L, 1% Tween 20, pH 2.2) twice for 10 min, washing twice in PBS for 10 min and washing in TBST twice for 5 min. Blots were then reblocked and reprobed.

2.3.6 Protein Kinase Assays

Purified, recombinant NEK 12 proteins (active and dead mutants) were adjusted to 100 µg/ml in 20 µl of KAB (50 mM MOPS, 20 mM MgCl₂, 10 mM DTT, 10 mM EGTA). A mastermix of KAM was prepared (50 µg/ml) containing 100 µM ATP (γ-³²P). Reactions were incubated at 30°C for 30 min with shaking and stopped by the addition of 40 µl 2x SDS-PAGE buffer and boiling for 3 min. Twenty microlitres of this was separated by SDS-PAGE, the gel was dried and placed into a cassette with a phosphor-storage screen (Molecular Dynamics) for 4 h to 24 h before visualising the screen in a Typhoon 9400 imager (GE Healthcare).

2.4 Trypanosome Cell Culture

2.4.1 Cell Lines and General Culturing Techniques.

The predominant cell line used in this study was *Trypanosoma brucei brucei* 2T1 which was gifted by Dr. David Horn. *T. b. brucei* 427 was occasionally used as a wild type control.

Trypanosomes were maintained in HMI-11 (HMI-9 (GIBCO), 10% v/v FCS (GIBCO 10270), Pen/Strep (SIGMA) (penicillin 20 U/ml, streptomycin 20 µg/ml)), at 37°C, 5% CO₂ in vented flasks. (Hirumi, H. Hirumi, 1989). When harvesting the cells for microscopy they were centrifuged at 800g for 10 mins. For other applications they were harvested at 1500g for 5 min (10 min for large volumes). Cells were maintained at densities below 2x10⁶ cells/ml.

Drug concentrations used for BSF *T. brucei* 2T1 (and derivatives): 5 µg/ml hygromycin B (Calbiochem), 2 µg/ml Puromycin, 2.5 µg/ml phleomycin (InvivoGen) and 10 µg/ml blasticidin (InvivoGen). Puromycin and phleomycin were used in the maintenance of untransfected 2T1s, switching to phleomycin and hygromycin after pTL transfection. Blasticidin was used to select for transfectants carrying tagging constructs.

When stabilating a cell line, 10 ml of mid-log cells were pelleted by centrifugation at 1500 g x 10 min and resuspended in 4ml HMI-11 + 10% glycerol v/v as a cryoprotectant. Aliquots of 1 ml were placed in cryovials (AlphaLaboratories, Feel the Seal 1 ml tubes) and placed at -80°C overnight then stored long-term under liquid nitrogen.

2.4.2 Genetic Modification of Trypanosomes

Bloodstream form parasites were grown to a density of 1x10⁶ cells/ml and 1x10⁷ cells were gathered for each transfection. Cells were collected by centrifugation at 1500 g for 10 minutes. Cells were resuspended in 100 µl of Amaxa T-Cell transfection buffer (Lonza). Five micrograms of linearised plasmid in 10 µl of water was added to this mix. Electroporation was conducted using an Amaxa Nucleofector 2 (Lonza) set to programme X-001. The transfected cells were placed into a 30 ml volume of HMI-11 and then diluted 1:20 and 1:40 in the case of RNAi plasmids and 1:10, 1:100 and 1:1000 in the case of vectors with unknown transfection efficiency. Selective drugs were added 6 hours after transfection. Clones were selected from the plates with the lowest number of positive wells.

2.4.3 Growth Curves

Thirty millilitres of HMI-11 was seeded to a density of 10^5 cells/ml and divided in 5 ml aliquots into a 6 well plate (Corning). Three wells were induced with tetracycline. Stock tetracycline (Sigma) was made in 70% ethanol at 1 mg/ml and added to cultures at 1 μ g/ml. Negative controls used the same volume of 70% ethanol. Cells were counted at 24 hour timepoints and cultures maintained below 1×10^6 cells/ml. Cell densities were calculated and dilution factors accounted for, allowing a cumulative growth curve to be generated.

2.4.4 Mouse Infections

ICR mice were inoculated with varying parasite numbers and RNAi induced at varying times depending on the severity of the in culture phenotype (detailed below). Inoculum volume was adjusted to 500 μ l in TDB. The concentration of the inoculum varied depending on the expected severity of the phenotype. For CRK3 *in vivo* RNAi mice were inoculated with 5×10^5 parasites and RNAi was induced 48 h post infection. For PDK1, ZFO and NEK 12.1 mice were inoculated with 5×10^4 parasites and RNAi induced 24h post infection. RNAi was induced by offering Doxycycline hyclate-laced (Sigma Aldrich D9891-5g), sugar water to mice (0.2 g/L plus 50 g/L sucrose). Mice with parasitaemia above 1×10^8 cells per ml were humanely culled. Parasitaemia was monitored using a haemocytometer. All animal work was conducted in accordance with Home Office regulations by licensed animal handlers.

2.4.5 Alamar Blue Screening for RNAi Phenotypes

Cell cultures were adjusted to 2×10^4 cells per ml and divided into two pools. One had RNAi induced by the addition of tetracycline (Sigma) in 70% ethanol to a final concentration of 1 μ g/ml the other treated with an equivalent volume of 70% ethanol, 200 μ l of each culture was then plated in triplicate into wells of a 96 well plate (Corning) and incubated for 48 h. Alamar blue (Resazurin 0.49 mM in PBS, 20 μ l) was added and the plates incubated for a further 24 h. The plate was then read at $\lambda_{\text{excitation}}$ 530 nm and $\lambda_{\text{emission}}$ 590 nm in an Envision Plate Reader (Perkin Elmer). The ratio of the values for induced wells over uninduced wells was calculated and used to derive a mean and standard deviation. A value of 1

indicated there was no effect of the RNAi on cell proliferation, <1 shows a decrease in proliferation and decreases further dependent on the severity of the growth defect.

2.5 Further Analyses

2.5.1 Flow Cytometry for DNA Content and EP Procyclin

Pellets were made from trypanosome cultures (containing 3×10^6 cells) and were washed in Trypanosome Dilution Buffer (TDB; 20 mM Na_2HPO_4 , 2 mM NaH_2PO_4 , 80 mM NaCl, 5 mM KCl, 1 mM MgSO_4 , 20 mM glucose, pH 7.4).

To determine DNA content of cells by flow cytometry cells were fixed in 70% methanol 30%:PBS overnight at 4 °C. Cells were washed in PBS then resuspended in PBS + 10 µg/ml propidium iodide, 10 µg/ml RNase A at 37°C for 45 min. Cell fluorescence and scattering parameters were measured on a FACS-Calibur Flow Cytometer (Becton Dickinson) using the FL-2A channel. Data analysis was performed using FlowJo software (Tree Star Inc.).

For the analysis of BSF to PCF differentiation the cells were fixed in 2% formaldehyde v/v, 0.05% glutaraldehyde v/v at 4°C for 1 hour, cells were stable in this buffer for several weeks but routinely analysed sooner. After washing in PBS once, the cells were labelled with FITC-conjugated anti-EP procyclin mouse monoclonal IgG₁ (CLP001F, clone TBRP1/247, Cedarlane Laboratories, Ontario) at a dilution of 1/100 in PBS for 1 hour on ice. Cells were washed once in PBS before being resuspended in 1 ml of PBS and filtered through a 40 µm Nitex mesh to remove aggregates. Cell fluorescence and scattering parameters were measured on a FACS-Calibur Flow Cytometer (Becton Dickinson) using the FL-1H channel. Data analysis was performed using FlowJo software (Tree Star Inc.).

2.5.2 Immunofluorescence for Epitope Tagged DFK

IFA protocols were provided by Dr Tansy Hammarton, University of Glasgow.

BSF cells (10^6) containing an endogenously tagged 12Myc-DFK fusin protein (cell line STL2378/STL2379) were pelleted at 800 x g for 10 min. Cells were washed in TDB and then resuspended in 100 µl PBS before 10 µl aliquots were spotted onto

poly-L-lysine coated microscope slides (Hendley, Essex) with wells divided by a PAP pen (Invitrogen). Cells were allowed to dry onto slides before being placed in 4% formaldehyde (Pierce) in PBS was placed on the wells for 5 minutes then washed twice in PBS for 5 min per wash. Wells were blocked using PBS/1%BSA for 1 hour. This step included 0.1% Triton X100 if a permeabilization step was required. Primary antibody was added in PBS for 1 h at room temperature in a humid box. The primary was removed and the wells washed twice with 50 μ l PBS. An appropriate secondary antibody was added at 1/200 dilution in PBS and incubated for 1 h in a dark humidified box. The secondary antibody was removed and the slide washed four times with 100 mM HEPES pH7.5 for 5 minutes before the addition of a small drop of VECTASHIELD + DAPI (Vector Laboratories). A cover slip was placed on the slide and sealed with nail varnish (Superdrug).

2.5.2.1 Microscopy

Routine microscopy was carried out on a Zeiss Axioskop UV microscope equipped with an Orca-ER camera (Hamamatsu) and Openlab software (Improvision). Images of two or more fluorescence channels and / or brightfield were merged using Photoshop CS (Adobe). Some images were obtained from an IN-Cell Analyzer 2 (GE Healthcare) and merged using ImageJ (NIH).

2.5.3 Cell Fractionation

Crude cell fractionation experiments were conducted by lysing pellets of 2×10^7 BSF trypanosomes under different buffer conditions. Pellets were washed in TDB. For hypotonic lysis cells were suspended in 500 μ l of 10 mM Tris-HCl, pH 7.4 and lysed with 4 x 15 s bursts of sonication in Fisher FB15047 water bath filled with ice-cold water. Detergent lysis was conducted by resuspending cell pellets in 500 μ l of 1% Triton X-100, 25 mM HEPES, 1 mM EDTA, pH 7.4 for 30 min on ice. The lysates were clarified by centrifugation at 14 000 g for 5 min. In both cases the supernatant was mixed with 125 μ l of 4x SDS-PAGE buffer with boiling for 5 min and the pellet was solubilised in 625 μ l of 1x SDS-PAGE buffer with boiling for 5 min. The fractions were analysed by SDS-PAGE where the equivalent of 1×10^6 cells was loaded per sample.

2.5.4 Transmission Immuno-Electron Microscopy

A pellet of 2×10^7 BSF trypanosomes expressing C-terminally 12xMyc tagged DFK (STL2378) was chemically fixed in 0.2% v/v glutaraldehyde, 2% formaldehyde w/v in PBS pH7.4 at room temperature for 10 min followed by 1 h on ice. Fixed pellets were transferred to the Integrated Microscopy Facility at the University of Glasgow, with procedures conducted by Ms Margaret Mullin. The pellets were prepared for transmission electron microscopy by the low temperature dehydration embedding protocol as follows.

The fixed cells were rinsed 3x5 min with PBS on ice, dehydrated through ethanol series as follows:

50% Ethanol - 2x5 min on ice.
 70% Ethanol - 3x10min at -40°C .
 90% Ethanol - 3x10 min at -40°C .
 100% Ethanol - 3x10 min at -40°C .

2:1 Absolute Ethanol:HM20 resin - 30 min at -40°C .
 1:1 Absolute Ethanol:HM20 resin - 30 min at -40°C .
 1:2 Absolute Ethanol:HM20 resin - 1 h at -40°C .
 Pure HM20 resin overnight at -40°C
 Several changes of Pure HM20 resin next day (x3) -40°C .

Cell pellets were embedded in gelatine capsules excluding air with cover slips and UV polymerised at -40°C for 24 h. 60-70 nm sections were cut using a Leica Ultracut UCT and picked up on 300mesh formvar coated Nickel grids. The sample sections were immuno labelled using Mouse Anti-Myc4A6 (Millipore) 1/10 or 1/100 for 1 h RT and Gold Anti-Mouse 1/20 for 1 h at room temperature. The samples were contrast stained with 0.5% aqueous Uranyl Acetate - 10mins and Reynolds Lead Citrate - 5mins before viewing on the TEM Leo 912AB OMEGA. Images were captured using Olympus iTEM soft imaging system.

2.5.5 Immunoprecipitation of DFK

Cultures of STL2378 BSF trypanosomes containing DFK-12Myc, and *T. b. brucei* 2T1 as a negative control, were scaled up to 500 ml to allow the harvesting of 5×10^8 - 1×10^9 parasites. These were pelleted by batch centrifugation at 1500 g for 10 min then combined into large pellets that were stored at -80°C until lysis. Pellets were lysed in 0.85 ml of lysis buffer (50 mM Tris, 50 mM NaCl, 1% v/v NP-

40, Complete EDTA-free protease inhibitors (Roche), Phosphatase Inhibitor Cocktail II (Merck) - which was also added to all further buffers) on ice for 10 min with pipetting and then clarified by centrifugation at 10, 000 g for 10 min at 4°C. These lysates were applied to columns from the c-Myc Tag IP/Co-IP Kit (Pierce) - containing mouse anti-cMyc conjugated to agarose beads- and incubated at 4°C for 1 h with rotation. The beads were washed three times using the TBS buffer supplied with the kit (25 mM Tris, 0.15 M NaCl, pH 7.2) then eluted using the included non-reducing elution buffer (0.3 M Tris•HCl, pH 6.8, 5% SDS, 50% glycerol).

3 Generation and Analysis of a Kinome-Wide RNAi Library in *Trypanosoma brucei*

3.1 Introduction

3.1.1 RNA Interference in *T. brucei*

Trypanosoma brucei was one of the first organisms in which RNA interference was discovered to function and this biological process has been exploited to allow functional genetics to probe this organism's biology. RNAi is a process where double stranded RNA is cleaved, processed to form siRNAs which are used to direct endonucleases to cleave target RNAs and is an essential part of post-transcriptional gene regulation and defence against dsRNA viruses and transposable elements.

In a landmark study, Fire and Mello (1998) demonstrated that double-stranded RNA could dramatically downregulate the expression of genes in *C. elegans* when present at only a few molecules per cell, at much lower concentrations than single stranded anti-sense RNA (Fire et al., 1998). Following shortly from this study it was discovered that the introduction of double stranded RNA targeting α -tubulin caused trypanosomes to round up and become "fat". Reduction of tubulin mRNA and protein expression resulting in disruption of microtubules, causing a cytokinetic defect, was found to be the cause (Ngô et al., 1998). Further analyses have allowed the identification of much of the RNAi machinery. Reviewed recently the process of RNAi is fairly well understood (Atayde et al., 2011; Kolev et al., 2011). *T. brucei* possesses two Dicer-like proteins (DCL1 and 2), one Argonaut (AGO) protein and two RNA interference factors (RIFs) that participate in generating the RNAi response. DCL1 is localised in the cytoplasm whereas DCL2 is nuclear. These each possess two RNase III motifs (RNase IIIa and b), bind dsRNA and cleave it into siRNA duplexes of 24-26nt (Shi et al., 2006). These are assisted into *Tb*AGO by *Tb*RIF4 and *Tb*RIF5 leading to the formation of the RNA-induced silencing complex (RISC) which separates the passenger strand from the duplex (Barnes et al., 2012). This leaves the guide strand in RISC allowing it to bind to the complementary target mRNA, which is then cleaved. Unlike the RNAi machinery in higher eukaryotes, *T. brucei* does not appear to possess RdRp (RNA dependent RNA polymerase), which allows the amplification

of the RNAi response (Kolev et al., 2011). Gene silencing by transfection of dsRNA is therefore transient and limited to one cell division cycle (Ngô et al., 1998). Thus in order to utilise RNAi as a tool for genetic analysis systems had to be developed to allow stable transfectants carrying an inducible dsRNA expression system to be generated. *Trypanosoma brucei* is not unique amongst the kinetoplastids in possessing the RNAi pathway; *T. congolense* and *Leishmania braziliensis* both possess functional RNAi and systems have been developed to exploit this for functional genetics. However, *T. brucei* is the most amenable organism to work with (Kolev et al., 2011).

3.1.2 RNAi as a Tool for Investigating Trypanosome Biology

The implications of being able to selectively inhibit expression of a protein by reducing its parent mRNA levels soon led to the development of RNAi as a tool for functional genetics. This led to a burgeoning of systems to perform RNAi in trypanosomes. The early work to generate stable transfectants carrying inducible systems for expressing exogenous or toxic genes was pioneered in the laboratories of Christine Clayton (Biebinger et al., 1997) and George Cross (Wirtz et al., 1999). These systems were then used to develop systems for expressing stem-loop and opposing promoter RNAi fragments (Shi et al., 2000; Wang et al., 2000). The differing systems have been compared and demonstrate that stem-loop RNAi, where the sense and anti-sense RNA strands are covalently bound by a linker sequence, demonstrate a more potent silencing of the target mRNA (Durand-Dubief et al., 2003). A large number of systems have been developed since in order to meet particular needs, for example some constructs work more effectively in PCF parasites than BSF cells due to the use of the PARP promoter to drive transcription. Vectors for use in BSF cells typically rely on the T7 promoter using cells modified to express T7 RNA polymerase (Shi et al., 2000; Alibu et al., 2005). Inducible RNAi constructs that contain deleterious inserts must be placed in a transcriptionally inactive region of the genome to ensure stable transfection. Vectors that randomly integrate into the ribosomal RNA (rRNA) spacer regions are plagued by the problem of inter-clone variability as *T. brucei* possesses 9 rRNA spacer regions in the haploid genome. Some of these are more transcriptionally-“silent” than others and the variation in the levels of transcription in these regions leads to variability in the expression of any cassette entering them (Biebinger et al., 1996). Both the problems of T7 RNA pol

toxicity and clonal variability were simultaneously resolved by tagging a single rRNA spacer locus and utilising expression vectors with an endogenous rRNA promoter (Alsford et al., 2005; Alsford and Horn, 2008). Alsford and co-workers modified BSF *T. b. brucei* Lister 427 to express the tet repressor, then an inducible, GFP expressing plasmid (hygR) was transfected in, and resulting cell lines were screened for robust expression that was not leaky in un-induced cells. After choosing a cell line, the reporter construct was replaced in a manner that removed the 5' region of the hygromycin resistance gene, this cell line was termed *T. b. brucei* 2T1. This could be re-constituted with appropriate plasmids containing the 5' of the *hyg* gene and an expression cassette under the control of a rRNA promoter and a tet operator. The tagged locus is located on chromosome 2a and actually resulted in increased transfection efficiency over untagged cell lines. Though the cell line only allows the analysis of BSF parasites the combination of an endogenous promoter, very low clonal variability and the efficacy of the stem-loop RNA produced make it a very reliable system for BSF RNAi cell lines. The negative aspect of the pRPa^{ISL} stem-loop RNAi constructs is the necessity to clone two inserts in opposing directions using different restriction enzymes. This can limit the target sequence selection and requires several sub-cloning steps. Our experience of the system was that of a low cloning efficiency, making this unsuitable for analysing large cohorts of genes.

3.1.3 RNAi Screens in *T. brucei*

Following the discovery of RNAi in trypanosomes and the generation of the early tools for conducting inducible RNAi, a number of studies worked to characterise individual genes of interest. As the tools for RNAi in trypanosomes became established, efforts began to intensify their use from a mix of candidate studies towards systematic interrogation of large numbers of genes or genome-wide screens.

Studies using RNAi on a candidate gene often then followed up initial phenotyping with highly detailed experiments to try and ascribe function to the target e.g. growth defects can be investigated for cell cycle effects. This can be intensive work and is not amenable to high-throughput screening. However, in 2006 the results of a multi-centre collaboration to perform RNAi studies against ORFs on chromosome 1 of were published (Subramaniam et al., 2006). The

TrypanoFAN project sought to expand the scale of reverse-genetic studies to at least a medium-throughput and targeted 197 of the 369 predicted ORFs on chromosome 1 (210 total after inclusion of other genes). VSG genes and pseudogenes were excluded, as well as those where RNAi could not be guaranteed to be specific (as assessed by the RNAi Software) (Redmond et al., 2003). The study was a huge task; analysing growth, motility, morphology, cell cycle, and the integrity of cellular processes and organelles of bloodstream form trypanosomes. Of the 210 knockdowns analysed 33% generated a significant phenotype (12% being lethal). The next largest systematic RNAi screen was undertaken in the laboratory of Christine Clayton; it targeted 37 genes (the products of which were predicted to contain RNA binding motifs) in both BSF and PFC trypanosomes and also attempted to localise some of the proteins (Wurst et al., 2009). Twenty-five of these genes generated RNAi phenotypes: 8 strong defects in BSF, 9 mild defects in BSF and 7 weak defects in PCF. The number of contributing workers in both the Wurst and Subramaniam studies highlights how labour intensive this type of candidate driven research is. Another, smaller screen targeted 30 protein kinase genes by RNAi, but appeared to be limited in its success and is discussed below (Mackey et al., 2011).

In stark contrast to the candidate approach is the genome-wide library approach, a technique which allows the response of a huge pool of parasites, carrying random RNAi target sequences, to a specific selection pressure or method. In 2002 the first ever example of a genome-wide RNAi screen (in any organism) was performed in the laboratory of Paul Englund (Morris et al., 2002). Genomic DNA was sheared by sonication into target sequences with a mean length of 600bp. These were blunt-end ligated into the pZJM vector and used to transfect procyclic form trypanosome. Enough trypanosomes were transfected to theoretically target the genome 5 times over. As individual, clonal cell lines can not be analysed using this technique a selection process is needed to generate meaningful data. In this study RNAi was induced and cells were passed over concanavalin A sepharose to bind cells that expressed EP procyclin. Cells that expressed RNAi targets against genes involved in controlling EP procyclin expression would not be bound to the column. Cells that did not bind to the column were further purified by FACS sorting and these were cloned out and the cultures expanded. The clones were then analysed by sequencing the RNAi

plasmid they carried allowing the identification of the target gene; in this case the clones that were identified had RNAi targets against the hexokinase gene showing that changes in the surface coat could be linked to changes in glycolysis.

This library was used in conjunction with other selection strategies to further probe trypanosome biology. One study used the compound tubercidin to select for drug resistance phenotypes in RNAi induced procyclic cultures and identified hexose transporters as being silenced (Drew et al., 2003). The authors also showed that silencing hexokinase also yielded tubercidin resistant parasites, implicating glycolysis as being the target pathway for tubercidin (and highlighted phosphoglycerate kinase as being the major target of the compound). This type of approach has since been used with other RNAi library systems to elucidate mechanisms of drug action and resistance of BSF trypanosomes to trypanocidal drugs (discussed below).

Using this type of library without selection, and instead cloning the parasites by dilution after transfection of the library, allowed individual cell lines to be screened for kDNA defects in another study, however this is a difficult way to maximise the returns from this type of library (Zhao et al., 2008). Another study utilising this type of approach identified cell cycle regulators including TOR1 and TOR2 kinases, but again highlighted the limitations of this type of approach (Monnerat et al., 2009). More recently two separate studies used RNAi libraries of differing construction to identify the transporters responsible for the uptake of melarsoprol and eflornithine and the activator of the nifurtimox pro-drug (nitroreductase enzyme) (Baker et al., 2011; Schumann Burkard et al., 2011).

In a paradigm shift of how genome-wide RNAi libraries are analysed, Alsford and colleagues used Illumina Sequencing to analyse in parallel the effect of the Englund genome -wide RNAi library, in various life-cycle stages and under various conditions (Alsford et al., 2011). The technique, known as RIT-seq (RNA interference target sequencing) was made possible due to the development of 2nd generation sequencing technologies and their subsequent ubiquity in modern molecular biology as well as developments to improve the transfection efficiency of *T. brucei*.

In order to achieve this increase in transfection efficiency a cell line termed Sce* was generated to allow transfection with the Englund RNAi library. This cell line inducibly expresses a homing endonuclease I-SceI from a cassette in the rRNA spacer region on chromosome 2, a region shown to give high levels of expression (Alsford et al., 2005; Glover and Horn, 2009). Tetracycline can be used to induce this enzyme several hours before transfection of the RNAi library. The I-SceI enzyme then induces double stranded breaks at sites flanking the expression cassette in the rRNA spacer region allowing it to be replaced by the RNAi cassettes. Transfected cells were drug selected and then RNAi was induced under various conditions (BSF, PCF and differentiating cells), DNA was then extracted and prepared for Illumina sequencing. Semi-specific PCR and size selection were used to enrich the sample for DNA containing part of the RNAi vector and an RNAi insert. Following sequencing these were mapped onto the genome sequence of *T. brucei* 927 and allowed the identification of “cold-spots”. These cold-spots correspond to genes that have been lost from the population, indicating that induction or RNAi for that sequence was lethal or reduced the parasites growth in the population and allowed identification of genes responsible for loss of fitness when silenced. The analysis was estimated to cover over 99% of non-redundant protein coding sequences in the genome (7435 ORFs) and to target each ORF five-fold. Significant loss of fitness phenotypes were reported for 1908 and 2724 CDS in BSF sample induced for 3 days and 6 days respectively, 1972 in PCF cells and 2677 in differentiating cells.

The RIT-seq approach has since been applied to other scenarios, which instead of looking for loss of fitness phenotypes look for genes which are not essential to the parasite and provide a selective advantage in response to certain pressures when silenced. In order to identify genes and pathways associated with drug action and resistance cultures of RNAi induced cell lines were exposed to lethal doses of various trypanocides. Surviving cells were analysed by RIT-seq and genes for which RNAi fragments were present were presented as “hot-spots” in the analysis, thus linking them to resistance to the individual trypanocidal drugs (Alsford et al., 2012). RIT-seq is a powerful technique but is not without its downsides. It still relies on having some form of assay that selects for a certain phenotype. The technique also relies on a vector containing opposing T7 promoters. These can be leaky, and T7 RNA polymerase itself is toxic to the

trypanosome. The sense and antisense RNA generated is not covalently bound so the penetrance of the RNAi effect may be lower than that from an equivalent stem-loop vector. Despite the power of genome-wide libraries it is in many cases advantageous to generate individual RNAi cell lines to study candidate genes in detail.

3.2 Research Aims

The aim of this part of the study was to improve the *T. brucei* 2T1 and pRPa^{ISL} system to allow RNAi of candidate genes to be performed in larger cohorts than previous techniques allowed. The aim was to make a system that would allow the screening of the kinome of *Trypanosoma brucei brucei*. The RNAi lines would be assessed for growth phenotypes and protein kinases considered to be good drug targets investigated further. Levels of target gene knockdown would be assessed by qPCR.

3.3 Results

3.3.1 Generation and Validation of the pTL System

Cloning of several RNAi constructs was undertaken using the pRPa^{ISL} (MCS1) vector as a backbone (Alsford and Horn, 2008). In an attempt to generate large numbers of RNAi constructs rapidly, a ligation independent cloning strategy (Clontech) was attempted, though no success was ever achieved using this system (Novy and Yaeger, 1996). Reverting to restriction enzyme cloning proved successful but very slow, so alternative methods were sought. It was noted that RNAi vectors for use in *Arabidopsis thaliana* had been developed that used Gateway technology to allow high-throughput cloning of constructs (Helliwell et al., 2002).

It was decided to modify the pRPa^{ISL} vector to contain repeated, inverted AttP sites from the Gateway system (Invitrogen), allowing one step-cloning of a single AttB-tagged PCR product into the vector, thus generating stem-loop RNAi constructs. The modification was synthesised as a cassette and inserted between the BamHI and XbaI sites of the vector backbone by Blue Heron (Bothwell, WA). This insert consisted of AttP1-ccdB-AttP2-Stuffer-AttP2-ccdB-AttP1 (Figure 3-1). Restriction enzyme sites for ClaI and StuI were inserted to allow digests of sense

and anti-sense inserts to assist in confirmation of correct integration of inserts. The stuffer-fragment was reduced from a 468bp fragment of the *lacZ* gene to a 150bp of the same gene to reduce the size of the synthesis and thus reduce the cost of the project. Predictions of RNA folding suggested this stuffer would still allow hairpin formation (Figure 3-2). Following recombination of an insert the RNAi fragment is flanked by 100 nt attL sites, these are also transcribed to dsRNA so the sequence was analysed to ensure it would not target part of the trypanosome genome. This was not the case and it was not anticipated that non-specific RNAi phenotypes would be generated.

Upon delivery of the plasmid a large batch was prepared and tested for purity by restriction enzyme digest (Figure 3-3). It was found that due to the inverted repeats and double complement of toxic *ccdB* genes the plasmid was unstable when grown at 37°C and was subsequently produced at 25°C. It was dubbed pGL2084 in our laboratory's plasmid database. Derivatives from this plasmid were termed pTLs (plasmid-Trypanosome Library), and oligos used to generate the RNAi inserts for a gene of interest were called TLO (Trypanosome Library Oligo). Collectively this was called the pTL System.

Figure 3-3 also shows the results of a restriction enzyme digest to validate a completed pTL plasmid targeting TbNEK 9 (Tb927.5.2820). pTL2 yields a simple digest pattern for exemplifying the mode of verification that was used for validating these plasmids. Restriction sites specific to the RNAi target sequence can allow additional verification.

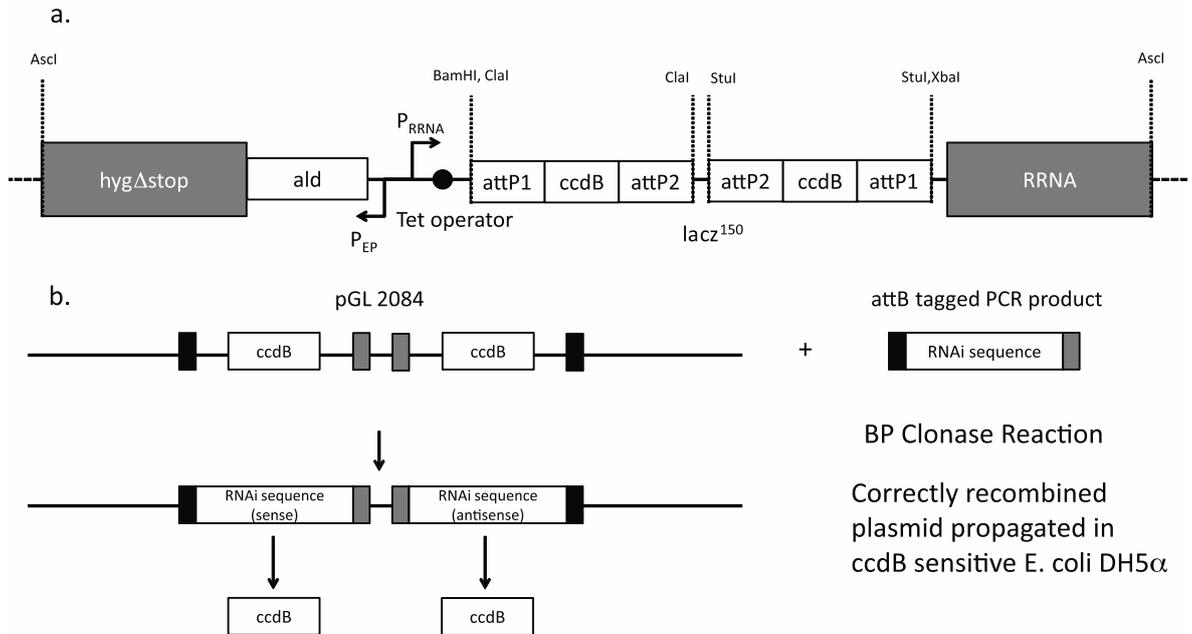


Figure 3-1: Schematic depicting Gateway cloning of inserts into pGL2084. Panel a. depicts main features of the plasmid required for cloning of inserts, integration into the *T. b. brucei* 2T1 cell line, and control of RNAi expression. After transfection the expression of the complete *hyg* gene is controlled by the EP procyclin promoter (P_{EP}) and the RNAi sequence by the ribosomal RNA promoter (P_{RRNA}). Panel b. depicts the cloning of an insert into the vector and the excision of the *ccdB* counter selectable marker.

Secondary structure: $\Delta G = -1587.1 \text{ kcal/mol}$

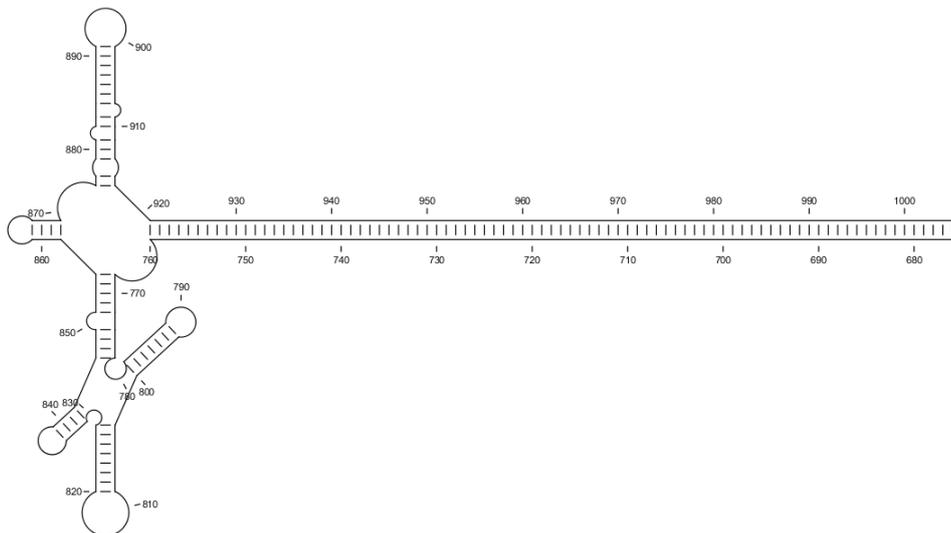
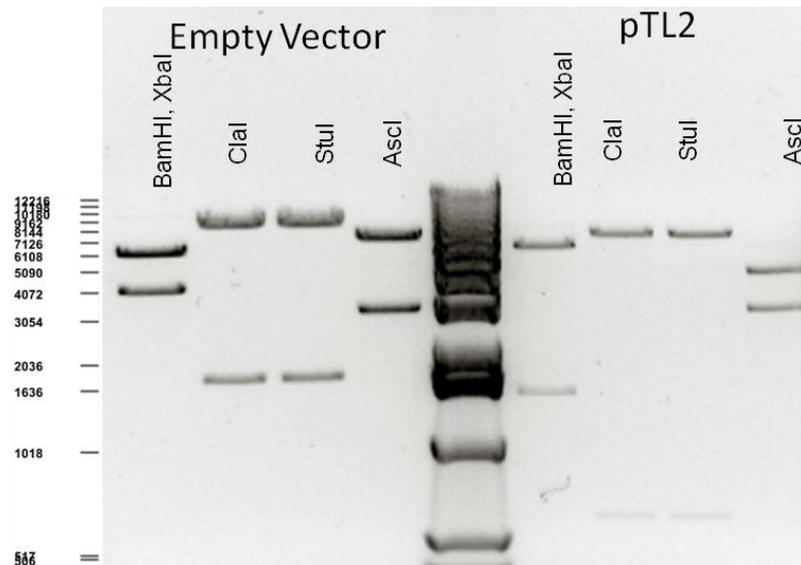


Figure 3-2: Prediction of hairpin folding from the redesigned, 150 nt stuffer fragment of pGL2084. RNA structure was predicted using CLC Genomic Workbench to assess whether the stuffer fragment would form a loop.



b.	Fragment Sizes bp			
	pGL2084		pTL vector	
	Insert	Backbone	Insert	Backbone
StuI	1577	6754	~600 (insert dependent)	5788
ClaI	1579	6756	~600 (insert dependent)	5790
Ascl	5729	2610	~3797 (insert dependent)	2610
BamHI+XbaI	3337	5002	1300-1600 (insert dependent)	5013

Figure 3-3: Validation of correct cloning by restriction endonucleases digest of pGL2084 and a correctly recombined plasmid (pTL2). Upper panel contains an example of the analysis for the pGL2084 vector and the pTL2 plasmid. Panel b. contains the expected fragment sizes for the various digest conditions.

3.3.2 Validation of the System

In order to validate this system a construct targeting TbCRK3 was generated and transfected into *T.b.brucei* 2T1. CRK3 is a positive control for an essential gene in trypanosomes and *Leishmania* and was used to prove the modified RNAi system could still function. Ablation of this gene by RNAi in trypanosomes rapidly results in cell cycle block at the G2/M boundary causing an increase in cells with multiple nuclei and kinetoplasts (Hassan et al., 2001; Tu and Wang, 2004). RNAi was induced in the cell line STL103 (containing the pTL21 plasmid to target CRK3) by addition of tetracycline and cell growth monitored over the following 72 hours. They were compared to an RNAi line (14006) generated with pGL1987 (this is derived from the original pRPa^{ISL} plasmid), which contains the same target sequence against CRK3 but was generated using conventional molecular cloning. It lacks the attL sequences found in pTL vectors and the stuffer

fragment is also longer, at 500 nt. However, the phenotype was indistinguishable between each cell line in terms of the growth defect (Figure 3-4). CRK3 mRNA levels were measured by qPCR and showed that both cell lines exhibited comparable reduction in target RNA (Figure 3-7).

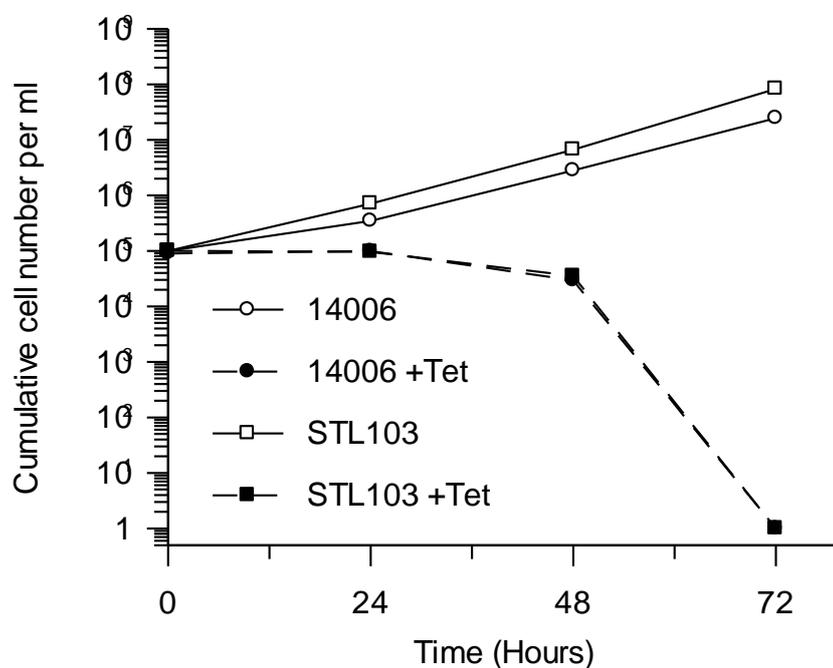


Figure 3-4: Growth curves of two cell lines targeting CRK3 using different RNAi systems. Cell line STL103 was generated using the pTL system and exhibits the same growth defect as cell line 14006 which was generated using pGL1875. The RNAi targeting sequence was the same in both plasmids. Cell line 14006 was created under an old system of stabilate numbering, STL represents the cell lines generated as part of the kinome-wide RNAi screen.

3.3.3 Generation of kinome-wide RNAi plasmid library

Following validation of the functionality of pGL2084, RNAi inserts were designed for all the protein kinases in the kinome by using RNAit to assess the CDS of each gene for suitable target regions. AttB sites were added to the 5' of each gene specific primer sequence and these oligos were used to generate PCR inserts. These were purified by gel electrophoresis followed by gel extraction. Inserts were cloned into pGL2084 and transformed into DH5 α MaxEfficiency cells due to their high transformation efficiency and their almost total inability to propagate ccdB-containing plasmids. After an initial period of optimisation, the process was carried out in batches. A QiaCube robot (Qiagen) allowed automation of purification steps leading to high quality DNA for Clonase reactions, such was the efficiency of these reactions that an average of 12 RNAi

constructs could be generated per week. The extremely high cloning efficiency of the Gateway system allowed the rapid generation of plasmids targeting all the genes in the kinome. Due to the inverted, repetitive nature of the DNA in this plasmid it proved difficult to sequence for confirmation of correct cloning. The *Ascl* digest used to linearise the plasmid appeared to give obvious differences between correct and incorrect plasmids so this digest was used to validate the plasmids to ensure proper recombination events.

Since implementing the pTL system we have generated over 250 plasmids for various gene families, including a set targeting 190 eukaryotic protein kinases (ePKs), atypical protein kinases (aPKs) and pseudokinase genes in *Trypanosoma brucei*. These plasmids were numbered PTL### and the library is named the PTL library (Plasmids, Trypanosome Library)

3.3.4 Generation of RNAi-cell lines

RNAi plasmids pTLs were purified from *E. coli* DH5 α and linearised using *Ascl*. Following ethanol precipitation this DNA was rehydrated and used to transfect bloodstream form *T.b.brucei* 2T1. After cloning by limiting dilution the lines were tested for loss of puromycin resistance, with puromycin sensitive clones being preferentially selected for analysis. These clones appeared to give more consistent and stable RNAi phenotypes. Cell lines were designated with an STL###, cryo-stabilated and screened for RNAi growth phenotypes. At least two independent clones were generated for each plasmid.

3.3.5 Alamar blue screening for growth defects under RNAi induction

In order to assess the growth phenotype of a limited number of RNAi clones, the gold standard is by counting cells in a Neubauer Improved Haemocytometer. This allows the cells to be counted accurately and sick or dead trypanosomes can easily be identified. However, this method is time consuming and is totally unsuitable for the medium-throughput size of the screen envisaged here. Other methods of counting cells were tested, including the Coulter Counter Z2, Millipore Sceptre and Invitrogen Countess but all were deemed unsuitable for various reasons. Methods of determining cell viability such as the MTT assay and alamar blue have been used to measure trypanosome viability and the alamar

blue assay has been well optimised for use in *T. brucei* 427 (Ráz et al., 1997). This assay was tested against various RNAi lines and was found to be an accurate way of detecting growth defects over a range of scenarios, from very weak growth defects (NEK 16) to very severe growth defects such as CRK3 and those falling in between (PDK1) (Figure 3-5).

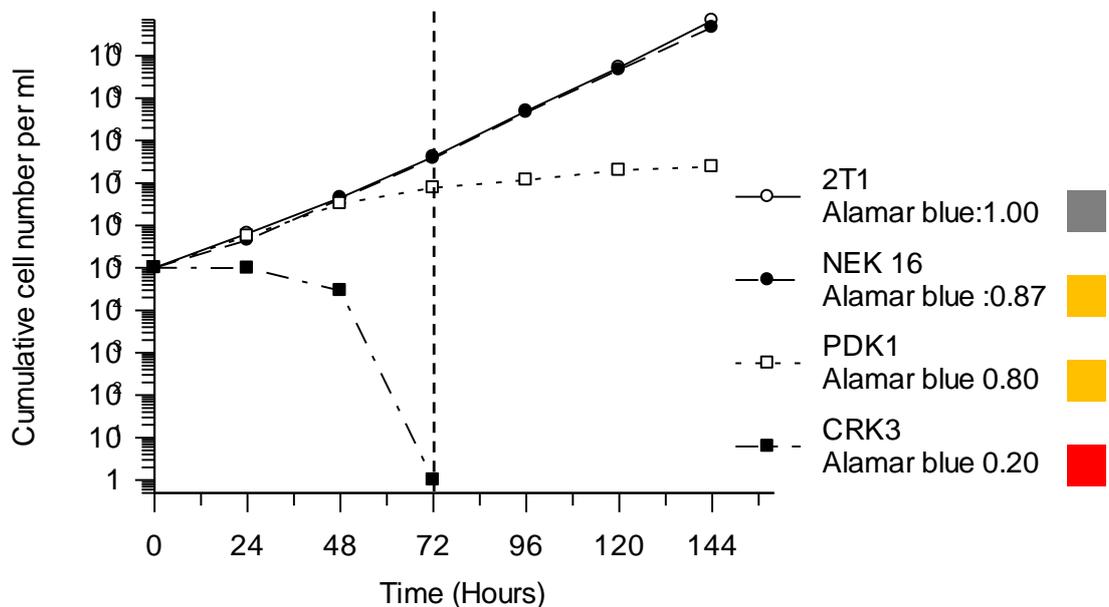


Figure 3-5: Growth curves for selected induced RNAi cell lines and a 2T1 control, with the corresponding values of the alamar blue assay which measures cell viability at 72h.

RNAi cell lines were assessed using the alamar blue cell viability assay. Induced and un-induced cultures were grown for 48h at which time alamar blue was added, 24 h after this addition the fluorescence of alamar blue at 590 nm was measured after excitation of 530 nm. The fluorescence value ratio for induced:un-induced was calculated and a mean of triplicate values generated. This value was thus calculated at 72 h after induction of RNA interference. This is the standard endpoint in an alamar blue assay and it was deemed to be suitable for the needs of this study as the aim was to detect strong phenotypes of essential protein kinases. It also allowed for rapid screening of the library to allow time for more detailed characterisation of essential/interesting phenotypes. Arbitrarily, cell lines showing a reduction in the ratio from 1 to 0.9 were classified as not showing a loss of fitness phenotype, those <0.9 - 0.5 were intermediate and those showing under 0.5 a severe loss of fitness phenotype.

Cell morphology was also observed during the course of the screening and very obvious phenotypes noted.

3.3.6 Results of a Kinome-Wide RNAi Screen

Following the transfection of the pTL RNAi library of plasmids to create individual cell lines targeting each protein kinase, two clones were assayed by the alamar blue method to detect growth defect phenotypes at 72 hours. This allowed the assessment of 190 genes for essentiality either individually or for a small subset of closely related genes, in double-knockdowns. Of the 190 genes targeted by RNAi constructs, 53 generated loss of fitness phenotypes detectable by alamar blue assay in the bloodstream form of the 2T1 strain of *Trypanosoma brucei brucei* (Figure 3-6). Twenty-six of these generated strong growth defects (<0.5 AB) (Table 3-1) and 27 generated milder loss of fitness phenotypes (0.5-0.9 AB) (Table 3-2). The remainder displayed no detectable growth defect by alamar blue (Table 3-3). Several of the identified genes resulted from double RNAi knockdowns but are known to be non-essential to trypanosomes (discussed below).

These data showed that all families of kinase in *Trypanosoma brucei* contained genes essential to the growth of bloodstream form cells, and that almost all contained genes that led to severe growth defects when ablated by RNAi. The protein kinases identified here are potential drug targets due to their essential role in BSF biology.

Table 3-1: Protein kinases determined to have a severe RNAi loss of fitness phenotype as determined by alamar blue ratio. * indicates an RNAi fragment that targets more than one gene. Rit-Seq Column denotes the phenotype detected by Alsford *et al.* 2011 – LOF; loss of fitness. Whether a clone lost puromycin resistance is denoted in the Puro R/S column. If a protein was detected as phosphorylated by Nett *et al.* 2009 this is denoted in the column BSF Phos.

Family	Name	Gene ID	RIT-Seq Screen	Clone 1 Puro R/S	Alamar Blue Ratio	Clone 2 Puro R/S	Alamar Blue Ratio	BSF Phos Nett <i>et al.</i>	Ref.
AGC/NDR	PK50	Tb927.10.4940	No LOF	Sensitive	0.21	Sensitive	0.86	Y	(Ma <i>et al.</i> , 2010)
aPK-PIKK/ATR	ATR	Tb11.01.6300	No LOF	Sensitive	0.43	Sensitive	0.82		
CK1/CK1*	CK1.1	Tb927.5.790	Day 3 BSF	Sensitive	0.16	Sensitive	0.16		(Urbaniak, 2009)
CK1/CK1*	CK1.2	Tb927.5.800	Day 6 BSF	Sensitive	0.16	Sensitive	0.16	Y	(Urbaniak, 2009)
CMGC		Tb11.02.3010	No LOF		0.30		0.37		
CMGC/CDK	CRK9	Tb927.2.4510	Day 3 BSF	Sensitive	0.17	Sensitive	0.18	Y	
CMGC/CDK	CRK12	Tb11.01.4130	Day 3 BSF	Sensitive	0.20	Sensitive	0.20		(Mackey <i>et al.</i> , 2011)
CMGC/CDK	CRK3	Tb927.10.4990	Day 3 BSF	Sensitive	0.32	Sensitive	0.27	Y	(Tu and Wang, 2004, 2005)
CMGC/CDK	CRK2	Tb927.7.7360	No LOF	Resistant	0.47	Sensitive	0.29	Y	(Tu and Wang, 2004, 2005)
CMGC/CLK		Tb11.01.4250	Day 3 BSF	Resistant	0.18	Sensitive	0.19		
CMGC/CLK		Tb11.01.4230	Day 3 BSF	Sensitive	0.21	Sensitive	0.26	Y	
CMGC/GSK	GSK3-short	Tb927.10.13780	Day 3 BSF	Sensitive	0.46	Resistant	0.61	Y	(Ojo <i>et al.</i> , 2008)
CMGC/MA PK		Tb927.10.5140	Day 3 BSF	Sensitive	0.21	Sensitive	0.22		
CMGC/MA PK	KFR1	Tb927.10.7780	No LOF	Resistant	0.42	Sensitive	0.55		(Hua and Wang, 1994, 1997)
CMGC/RCK		Tb927.3.690	Day 6 BSF	Resistant	0.23	Resistant	0.18		
NEK*	NEK12.2	Tb927.4.5310	Day 6 BSF	Sensitive	0.19	Sensitive	Growth Curves	Y	
NEK*	NEK12.1	Tb927.8.7110	No LOF	Sensitive	0.19	Sensitive	Growth Curves	Y	

Orphan	PK6	Tb09.211.2260	Day 3 BSF	Sensitive	0.24	Sensitive	0.23	Y
Other/AUR	AUK1	Tb11.01.0330	Day 3 BSF	Sensitive	0.19	Resistant	0.19	
Other/AUR		Tb09.160.0570	Day 6 BSF	Sensitive	0.35	Resistant	0.46	
Other/TLK*	TLK2	Tb927.8.7220	Day 6 BSF	Sensitive	0.21	Resistant	0.22	(Li et al., 2007b)
Other/TLK*	TLK1	Tb927.4.5180	No LOF	Sensitive	0.21	Resistant	0.22	(Li et al., 2007b)
STE		Tb927.10.2040	No LOF	Sensitive	0.32	Sensitive	0.26	
STE/STE11		Tb11.46.0003	No LOF	Sensitive	0.20	Resistant	0.20	
STE/STE20	SLK2	Tb09.211.3820	No LOF	Resistant	0.32	Resistant	0.78	Y
STE/STE20	SLK1	Tb927.8.5730	No LOF	Resistant	0.46	Sensitive	0.21	Y

Table 3-2: Protein kinases detected to have an RNAi loss of fitness phenotype as determined by an alamar blue assay ration between 0.5 and 0.9. * indicates a RNAi targeted more than 1 gene. # indicates a morphological phenotype was detected despite no change in alamar blue value. Rit-Seq Column denotes the phenotype detected by Alsford *et al.* 2011 – LOF; loss of fitness. Whether a clone lost puromycin resistance is denoted in the Puro R/S column. If a protein was detected as phosphorylated by Nett *et al.* 2009 this is denoted in the column BSF Phos.

Family	Name	Gene ID	RIT-Seq Screen	Clone 1 Puro R/S	Alamar Blue Ratio	Clone 2 Puro R/S	Alamar Blue Ratio	BSF Phosphorylation Nett <i>et al.</i>	Ref.
AGC		Tb927.3.2440	Day 6 BSF	Sensitive	0.69	Sensitive	0.62	Y	
AGC	TbPDK1	Tb09.160.3480	No LOF	Sensitive	0.80	Sensitive	Two identical growth curves		
AGC/NDR	PK53	Tb927.7.5770	No LOF	Sensitive	0.84	Sensitive	1.00	Y	(Ma <i>et al.</i> , 2010)
AGC/PKA*	PKAC1	Tb09.211.2410	Day 3 BSF	Sensitive	0.53	Sensitive	0.50	Y	
AGC/PKA*	PKAC2	Tb09.211.2360	Day 6 BSF	Sensitive	0.53	Sensitive	0.50		
aPK-ABC1		Tb927.10.9900	No LOF	Resistant	0.70	Sensitive	0.88		
aPK-PIKK/FRAP	TOR 1	Tb927.10.8420	No LOF	Sensitive	0.56	Sensitive	0.62		
aPK-PIKK/FRAP	Tor 4	Tb927.1.1930	Day 3 BSF	Sensitive	0.71	Sensitive	0.92	Y	
aPK-RIO	RIO2	Tb927.6.2840	No LOF	Resistant	0.72	Resistant	0.70		
aPK-RIO	RIO1	Tb927.3.5400	No LOF	Sensitive	0.84	Sensitive	0.98		
CAMK		Tb927.7.6220	No LOF	Sensitive	0.77	Sensitive	0.84		
CAMK	LDK	Tb11.01.0670	No LOF	Sensitive	0.81	Sensitive	0.78		(Flaspohler <i>et al.</i> , 2010)
CAMK		Tb927.7.2750	No LOF	Sensitive	0.83	Sensitive	0.75		
CAMK/CAM KL		Tb927.8.870	No LOF	Sensitive	0.87	Sensitive	0.87		
CK1/CK1	CK1	Tb927.3.1630	No LOF	Sensitive	0.87	Sensitive	1.02		

CMGC/CDK	CRK6	Tb11.47.0031	No LOF	Sensitive	0.76	Sensitive	0.76		(Tu and Wang, 2004, 2005)
CMGC/CDK	CRK1	Tb927.10.1070	Day 3 BSF	Resistant	0.81	Resistant	0.76		(Tu and Wang, 2004, 2005)
CMGC/DYRK		Tb927.10.9600	No LOF	Resistant	0.85	Resistant	0.81		
CMGC/SRPK		Tb927.6.4970	No LOF	Sensitive	0.68	Resistant	0.78		
Other/CAM KK		Tb927.10.15300	No LOF	Resistant	0.79	Resistant	0.77	Y	
Other/CK2	CK2A1	Tb09.211.4890	No LOF	Sensitive	0.63	Sensitive	0.65	Y	
Other/NEK	NEK16	Tb927.10.14420	Day 3 BSF	Sensitive	0.88	Sensitive	0.86		
Other/ULK		Tb11.02.2050	Day 3 BSF	Sensitive	0.70	Resistant	0.87		
pseudo Other/NAK		Tb09.160.4770	Day 3 BSF	Sensitive	0.57	Sensitive	0.71	Y	
pseudo Other/ULK		Tb11.01.0400	No LOF	Resistant	0.83	Resistant	0.77		
Pseudo- Orphan		Tb927.7.3210	Day 6 BSF	Sensitive	0.89	Resistant	0.75		
STE/STE11		Tb927.8.1100	No LOF	Sensitive	0.84	Sensitive	0.83		

Table 3-3: Protein kinases for which no loss of fitness phenotype was detected by alamar blue assay. Those marked with an asterisk had a visible morphological defect. Rit-Seq Column denotes the phenotype detected by Alford *et al.* 2011 – LOF; loss of fitness. Whether a clone lost puromycin resistance is denoted in the Puro R/S column. If a protein was detected as phosphorylated by Nett *et al.* 2009 this is denoted in the column BSF Phos.

Family	Name	Gene ID	RIT-Seq Screen	Clone 1 Puro R/S	Alamar Blue Ratio	Clone 2 Puro R/S	Alamar Blue Ratio	BSF Phosphorylation Nett <i>et al.</i>	Ref.
AGC	ZFK	Tb11.01.1030	No LOF	Resistant	1.00	Sensitive	0.99	Y	(Vassella et al., 2001)
AGC	RAC	Tb927.6.2250	No LOF	Resistant	1.01	Resistant	0.99		
AGC	putative PKB	Tb927.6.2450	Day 3 BSF	Sensitive	1.01	Sensitive	0.98		
AGC		Tb927.4.4590	No LOF	Sensitive	1.04	Sensitive	1.01		
AGC/PKA	PKAC3	Tb927.10.13010	No LOF	Sensitive	1.00	Sensitive	0.99	Y	
AGC/RSK		Tb927.3.2690	No LOF	Sensitive	0.97	Sensitive	1.01		
aPK		Tb927.1.1000	No LOF	Sensitive	1.02	Sensitive	1.03		
aPK		Tb927.8.7450		Sensitive	1.04	Sensitive	1.02		
aPK-A6		Tb927.3.5180	No LOF	Sensitive	0.97	Sensitive	1.05		
aPK-ABC1		Tb927.10.8730	No LOF	Sensitive	0.92	Sensitive	0.93		
aPK-ABC1		Tb927.8.4000	No LOF	Sensitive	1.01	Sensitive	0.98		
aPK-ABC1		Tb927.5.3420	No LOF	Sensitive	1.02	Sensitive	1.02		
aPK-ABC1		Tb927.10.11940	No LOF	Sensitive	1.04	Resistant	1.02		
aPK-Bud32		Tb11.03.0290	No LOF	Sensitive	0.95	Sensitive	1.00		
aPK-PDK		Tb11.02.2390	No LOF	Sensitive	1.01	Sensitive	1.02		
aPK-PDK		Tb927.4.840	No LOF	Sensitive	1.03	Sensitive	1.03		
aPK-PIKK/ATM	ATM	Tb927.2.2260	No LOF	Sensitive	0.99	Sensitive	0.98		
aPK-PIKK/FRAP	TOR 2	Tb927.4.420	No LOF	Sensitive	1.00	Sensitive	0.92		
aPK-	TOR-	Tb927.4.800	No LOF	Sensitive	1.00	Sensitive	1.03	Y	

PIKK/FRAP		LIKE1						
aPKS		Tb927.4.4970	No LOF	Resistant	1.01	ND	ND	
CAMK		Tb09.160.0500	Day 6 BSF	Sensitive	0.98	Resistant	0.95	
CAMK		Tb927.5.4430	No LOF	Sensitive	0.98	Sensitive	1.00	
CAMK		Tb927.4.3770	No LOF	Sensitive	0.98	Resistant	1.03	
CAMK		Tb927.10.14770	No LOF	Sensitive	1.04	Resistant	1.11	
CAMK/CAM KL		Tb927.2.1820	Day 6 BSF	Sensitive	0.92	Sensitive	1.07	Y
CAMK/CAM KL		Tb927.10.3900	No LOF	Resistant	0.97	Sensitive	1.09	Y
CAMK/CAM KL		Tb927.10.5310	No LOF	Sensitive	1.01	Resistant	1.01	
CAMK/CAM KL		Tb927.10.13480	Day 6 BSF	Sensitive	1.05	Sensitive	1.06	Y
CAMK/CAM KL		Tb927.10.13490	Day 6 BSF	Sensitive	1.05	Sensitive	1.06	
CAMK/CAM KL		Tb927.3.4560	No LOF	Sensitive	1.13	Sensitive	1.14	Y
CK1/CK1		Tb927.10.2390	Day 3 BSF	Sensitive	1.01	Resistant	0.98	
CK1/TTBK		Tb927.4.1700	No LOF	Sensitive	1.06	Resistant	1.05	
CMGC		Tb927.10.10870	No LOF	Sensitive	1.01	Sensitive	1.01	
CMGC/CDK	CRK4	(Tu and Wang, 2004, 2005)	No LOF	Sensitive	0.99	Sensitive	1.01	
CMGC/CDK	CRK11	Tb927.6.3110	No LOF	Resistant	0.99	Resistant	0.98	
CMGC/CDK	CRK8	Tb11.02.5010	No LOF	Sensitive	1.00	Sensitive	1.01	
CMGC/CDK	CRK7	Tb927.7.1900	Day 6 BSF	Sensitive	1.01	Sensitive	1.00	
CMGC/CDK	CRK10	Tb927.3.4670	Day 6 BSF	Sensitive	1.03	Sensitive	1.03	
CMGC/CDKL		Tb11.01.0380	No LOF	Resistant	1.01	Sensitive	0.95	
CMGC/CDKL	ECK1	Tb11.01.8550	Day 6 BSF	Sensitive	1.02	Sensitive	1.00	Y (Ellis et al., 2004)
CMGC/CLK		Tb927.2.4200	No LOF	Sensitive	1.00	Sensitive	1.01	

CMGC/CLK		Tb927.3.1610	No LOF	Resistant	1.00	Sensitive	1.04		
CMGC/DYRK		Tb927.7.3880	Day 6 BSF	Sensitive	0.97	Resistant	0.97		
CMGC/DYRK		Tb927.7.4090	No LOF	Sensitive	0.99	Sensitive	0.98		
CMGC/DYRK	TbDYRK 2	Tb11.02.0640	Day 3 BSF	Sensitive	1.01	Sensitive	1.01	Y	
CMGC/DYRK		Tb927.10.15020	No LOF	Sensitive	1.01	Sensitive	1.01	Y	
CMGC/DYRK		Tb927.5.1650	No LOF	Resistant	1.02	Sensitive	1.01		
CMGC/DYRK	PK4	Tb927.10.350	No LOF	Resistant	1.03	Resistant	1.02	Y	
CMGC/GSK	GSK3- long	Tb927.7.2420	No LOF	Sensitive	0.92	Sensitive	1.07		(Ojo et al., 2008)
CMGC/MAP K	mapk3	Tb927.8.3550	No LOF	Sensitive	0.95	Resistant	1.01		
CMGC/MAP K	mapk	Tb927.8.3770	No LOF	Resistant	0.97	Resistant	0.98	Y	
CMGC/MAP K		Tb927.10.12040	No LOF	Resistant	0.97	Resistant	1.00		
CMGC/MAP K	MAPK4	Tb927.6.1780	No LOF	Resistant	1.00	Resistant	0.95	Y	(Güttinger et al., 2007)
CMGC/MAP K	MAPK2	Tb927.10.16030	No LOF	Resistant	1.00	Sensitive	0.97	Y	(Müller et al., 2002a)
CMGC/MAP K	MAPK5	Tb927.6.4220	Day 6 BSF	Sensitive	1.01	Sensitive	1.01	Y	(Domenicali Pfister et al., 2006a)
CMGC/MAP K		Tb927.10.3230	No LOF	Sensitive	1.02	Resistant	1.03	Y	
CMGC/RCK	MAPK9	Tb927.10.14800	No LOF	Sensitive	1.01	Sensitive	1.02	Y	
CMGC/RCK- MAPK	MOK	Tb09.211.0960	No LOF	Sensitive	0.93	Sensitive	0.94	Y	
CMGC/SRPK		Tb927.7.960	No LOF	Sensitive	1.00	Sensitive	1.01		
NEK	NEK6	Tb927.3.3080	No LOF	Sensitive	0.94	Sensitive	0.99		
NEK	NEK-19	Tb927.10.1380	No LOF	Resistant	0.98	Sensitive	0.98		

NEK	NEK-22	Tb927.2.2120	No LOF	Resistant	0.98	Resistant	1.01	
NEK	NEK-10	Tb927.6.3430	No LOF	Sensitive	0.99	Resistant	0.95	
NEK	NEK-21	Tb11.01.6650	No LOF	Sensitive	0.99	Sensitive	0.99	
NEK	NEK-15	Tb09.160.0450	No LOF	Sensitive	0.99	Sensitive	1.01	
NEK	NEK-9	Tb927.5.2820	No LOF	Sensitive	1.00	Sensitive	1.01	
NEK	NRKA	Tb927.4.5390	No LOF	Resistant	1.00	Resistant	0.99	
NEK	NRKB	Tb927.8.6930	No LOF	Resistant	1.00	Resistant	0.99	
NEK	NEK20	Tb11.01.2900	Day 6 BSF	Resistant	1.01	Sensitive	1.01	
NEK	NRKC	Tb927.10.460	No LOF	Resistant	1.02	ND	ND	Y
NEK	NEK11	Tb927.7.3580	No LOF	Sensitive	1.03	Sensitive	1.01	
NEK	NEK7	Tb927.3.3190	No LOF	Sensitive	1.03	Sensitive	1.04	
NEK*	NEK-14	Tb927.8.1670	No LOF	Sensitive	0.96	Resistant	0.97	
NEK*	NEK-14	Tb927.8.1690	No LOF	Sensitive	0.96	Resistant	0.97	
NEK*	NEK-17	Tb927.10.5950	No LOF	Resistant	0.99	Resistant	1.02	Y
NEK*		Tb927.10.5940	No LOF	Resistant	0.99	Resistant	1.02	Y
NEK*		Tb927.10.5930	No LOF	Resistant	0.99	Resistant	1.02	
Orphan		Tb927.10.3340	No LOF	Sensitive	0.99		1.01	Y
Orphan		Tb927.3.3290	No LOF	Sensitive	1.00	ND-growth curves I think	n/a	
Orphan		Tb11.03.0340	No LOF	Sensitive	1.00	Sensitive	1.01	Y
Orphan		Tb11.01.2290	Day 3 BSF	Sensitive	1.01	Sensitive	0.98	
Orphan		Tb11.02.3640	No LOF	Resistant	1.01	Resistant	1.03	
Orphan	KinER5	Tb927.5.3150	No LOF	Sensitive	1.01	Sensitive	1.01	
Orphan		Tb09.160.0930	No LOF	Sensitive	1.02	Sensitive	1.05	
Orphan		Tb927.8.5500	No LOF	Sensitive	1.02	Sensitive	1.00	
Orphan		Tb927.10.1460	No LOF	Resistant	1.02	Sensitive	1.01	
Orphan		Tb927.5.3160	No LOF	Sensitive	1.03	ND	n/a	
Orphan		Tb11.02.4530	No LOF	Sensitive	1.03	Sensitive	1.02	

Orphan		Tb927.3.1850	No LOF	Sensitive	1.03	Sensitive	1.01	Y	
Orphan		Tb927.3.1570	No LOF	Resistant	1.03	Resistant	1.07		
Orphan		Tb927.10.15880	No LOF	Resistant	1.04	Resistant	1.01		
Orphan		Tb927.10.16160	No LOF	Sensitive	1.05	Sensitive	1.04		
Orphan	FlaK	Tb927.2.2720	No LOF	Sensitive	1.05	Sensitive	1.02		
Orphan		Tb927.1.3130	No LOF	Sensitive	1.06	Sensitive	1.01		
Orphan		Tb927.8.4260	No LOF	Resistant	1.06	NA	NA		
Orphan		Tb09.160.1090	No LOF	Resistant	1.08	Resistant	1.07		
Other/AUR		Tb927.3.3920	Day 3 BSF	Sensitive	1.00	Sensitive	1.01		
Other/CAM KK		Tb11.02.4860	No LOF	Sensitive	1.02	Resistant	1.04	Y	
Other/CAM KK		Tb11.02.4830	Day 3 BSF	Resistant	1.04	Resistant	1.06		
Other/CAM KK		Tb927.8.6490	No LOF	Sensitive	1.06	Resistant	1.05		
Other/CK2	CK2A2	Tb927.2.2430	No LOF	Sensitive	1.04	Sensitive	1.01		
Other/PEK	TbEIFK1	Tb11.02.5050	No LOF	Sensitive	1.01	Sensitive	1.01		
Other/PEK	TbEIFK3	Tb927.6.2980	No LOF	Sensitive	1.04	Resistant	1.03		
Other/PEK	EIF2K2	Tb927.4.2500	No LOF	Sensitive	1.05	Sensitive	1.03		
Other/PLK	PLK	Tb927.7.6310	Day 3 BSF	Sensitive	0.93	Sensitive	0.93	Y	(Li et al.; Hammarton et al., 2007; de Graffenried et al., 2008; Umeyama and Wang, 2008; Sun and Wang, 2011; Ikeda, 2012)
Other/PLK		Tb927.6.5100	No LOF	Sensitive	1.01	Sensitive	0.97		
Other/VPS1		Tb11.01.0930	No LOF	Sensitive	1.03	Resistant	0.91		

5								
Other/WEE	Wee1	Tb927.4.3420	No LOF	Sensitive	0.98	Resistant	0.99	
pseudo CAMK		Tb927.2.5230	Day 3 BSF	Sensitive	1.03	Resistant	1.02	
Pseudo-CAMK/CAM KL		Tb927.10.1940	No LOF	Sensitive	1.03	Resistant	1.02	
Pseudo-Orphan		Tb11.01.1050	No LOF	Resistant	0.91	Resistant	1.00	Y
Pseudo-Orphan		Tb927.5.3320	No LOF	Sensitive	1.00	Resistant	0.97	
Pseudo-Orphan		Tb927.7.3190	No LOF	Sensitive	1.01	Sensitive	1.00	
Pseudo-Orphan		Tb927.4.2460	No LOF	Resistant	1.01	Resistant	1.00	
Pseudo-Orphan		Tb09.211.1570	No LOF	Resistant	1.02	Resistant	1.01	
Pseudo-Orphan		Tb09.160.0990	No LOF	Resistant	1.04	Resistant	1.00	
Pseudo-Orphan		Tb11.01.6100	No LOF	Resistant	1.07	ND	n/a	
STE		Tb927.8.5950	No LOF	Sensitive	0.97	Resistant	0.96	Y
STE		Tb927.3.2060	No LOF	Sensitive	0.97	Sensitive	1.00	Y
STE		Tb927.1.1530	No LOF	Sensitive	0.99	Sensitive	1.04	Y
STE	kinER4	Tb927.3.5650	No LOF	Sensitive	1.00	Resistant	0.99	
STE		Tb11.01.5880	No LOF	Sensitive	1.00	Sensitive	1.03	Y
STE	MKK1	Tb927.3.4860	No LOF	Resistant	1.02	Resistant	1.00	(Jensen et al., 2011)
STE/STE11		Tb927.7.6680	No LOF	Sensitive	0.96	Sensitive	0.95	
STE/STE11	KinER3	Tb09.211.3410	No LOF	Sensitive	0.98	Sensitive	0.97	
STE/STE11		Tb927.6.2030	No LOF	Sensitive	0.98	Sensitive	0.97	Y

STE/STE11	Forkhead Kinase	Tb927.7.5220	No LOF	Resistant	0.99	Sensitive	1.14	
STE/STE11		Tb927.10.10350	No LOF	Sensitive	0.99	Resistant	1.01	
STE/STE11		Tb927.8.6810	No LOF	Sensitive	1.00	Sensitive	1.08	
STE/STE11		Tb927.10.14300	Day 6 BSF	Sensitive	1.00	Sensitive	1.00	Y
STE/STE11	KinER2	Tb09.160.1780	No LOF	Resistant	1.00	Resistant	1.01	
STE/STE11		Tb927.10.14780	Day 6 BSF	Sensitive	1.03	Sensitive	1.00	
STE/STE11		Tb927.10.1910	No LOF	Sensitive	1.04	Sensitive	1.01	
STE/STE11		Tb927.7.3650	No LOF	Sensitive	1.09	Sensitive	1.00	
STE/STE11#	DFK/KinER1	Tb11.01.5650	No LOF	Sensitive	0.96	Sensitive	Growth Curves	
STE/STE7		Tb927.8.1780	No LOF	Resistant	1.00	Sensitive	1.00	
STE/STE7	MKK5	Tb927.10.5270	No LOF	Resistant	1.02	Sensitive	1.02	(Jensen et al., 2011)

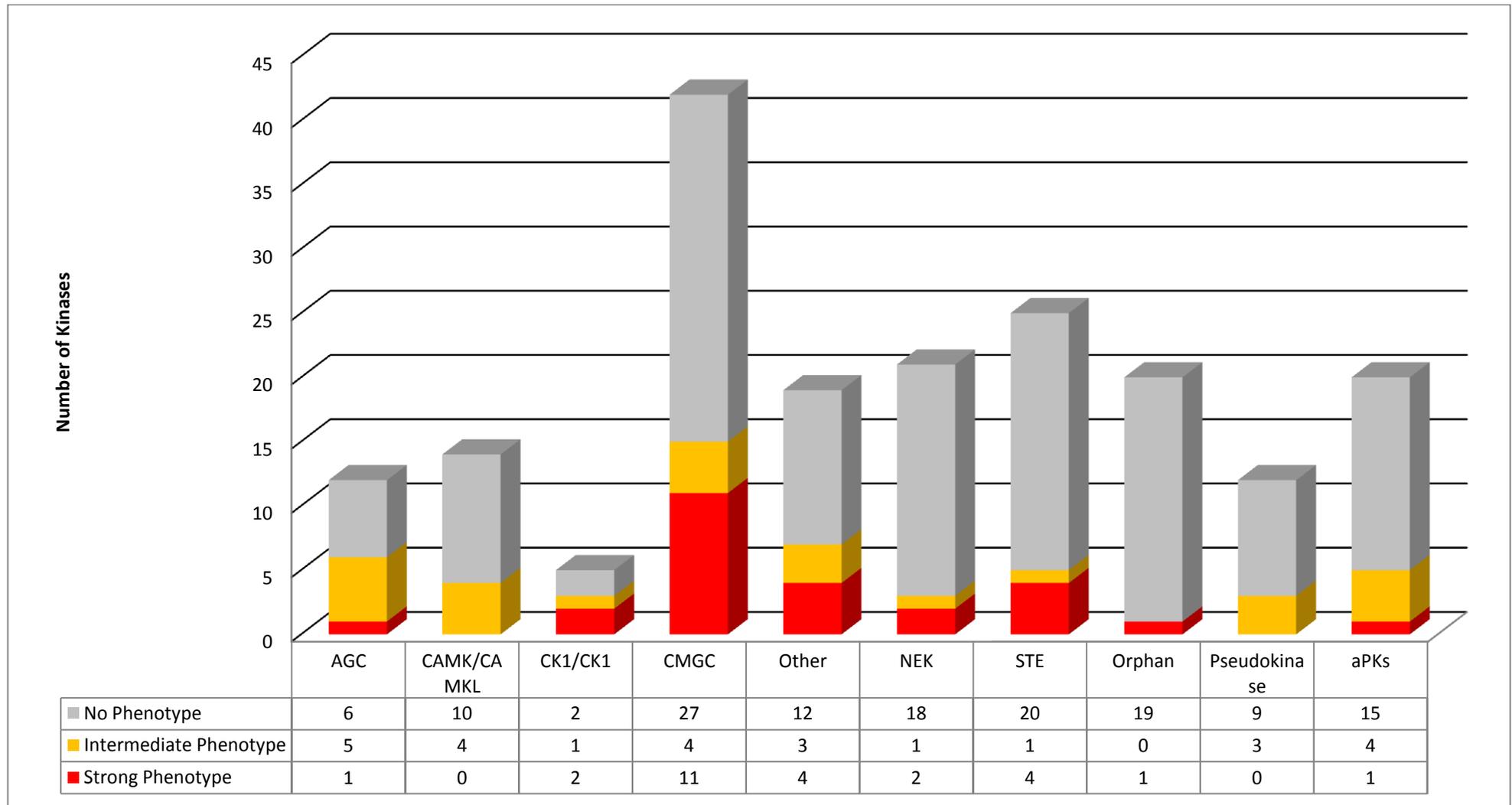


Figure 3-6: Results of a kinome-wide RNAi screen. The kinome has been divided into the major families and broken down by RNAi alamar blue phenotype, Red bars depict strong RNAi phenotypes (alamar blue value <0.5) and Orange bars denote moderate phenotype (0.5-0.9 alamar blue value).

Some RNAi constructs targeted multiple genes (for example TLK1/TLK2 and CK1.1/CK1.2), due to the similarity of the CDS regions used to design the RNAi inserts (listed in Table 3-4). Several of these have been the subject of more detailed, published studies. The RNAi target sequences used in this study were designed against the CDS of the genes and did not include UTRs, which may have limited the level of specificity achieved. TLK2 was shown not to be essential in PCF trypanosomes while TLK1 was essential, specificity being achieved by targeting the 3' UTR (Li et al., 2007b). TLK1 was later shown to be essential for BSF forms by RNAi, but no work to characterise TLK2 was shown for BSF parasites (Li et al., 2008). CK1.1 and CK1.2 have unambiguous evidence to define their essentiality in BSF parasites as CK1.1 has been knocked-out in BSF parasites with no detectable growth defect in culture. CK1.2 was shown to be essential in BSF by RNAi which targeted a region of the 5' UTR and the CDS unique to CK1.2. Other genes targeted by dual RNAi have suggestions of essentiality by RIT-Seq analysis, though during this analysis redundant reads were discarded which may impact on the detection of loss of fitness phenotypes on highly similar genes (Alsford et al., 2011). The two CLKs (Tb11.01.4230 and Tb11.01.4250) are dually targeted by two plasmids (pTL81/82); this seemed to be due to an error in RNAi where it did not identify them as paralogues. These have been re-designed to produce individual RNAi plasmids but remain to be tested.

Table 3-4: List of RNAi plasmids targeting multiple genes. The pTL screen phenotypes are depicted by colour (red: strong, orange: intermediate, grey: not detected) and the published phenotype and reference included for comparison. Note that by the RIT-Seq technique the phenotype is difficult to determine due to high sequence homology between the genes in this table.

Plasmid	Family	Genes Targeted	Name	Published Phenotype	Reference
pTL 46	AGC/PKA	Tb09.211.2410	PKAC1	LOF phenotype BSF by RIT-Seq	(Alsford et al., 2011)
	AGC/PKA	Tb09.211.2360	PKAC2	No LOF Phenotype by RIT-Seq	
pTL 204	CK1/CK1	Tb927.5.790	CK1.1	Not essential by double knockout	(Urbaniak, 2009)
	CK1/CK1	Tb927.5.800	CK1.2	Essential by RNAi	
pTL 167	Other/TLK	Tb927.8.7220	TLK2	Not essential by RNAi (PCF)	(Li et al., 2007b)
	Other/TLK	Tb927.4.5180	TLK1	Essential by RNAi (BSF/PCF)	
pGL 2064	Other/NEK	Tb927.4.5310	NEK12.2	LOF phenotype BSF by RIT-Seq	(Alsford et al., 2011)
	Other/NEK	Tb927.8.7110	NEK12.1	No LOF Phenotypes by RIT-Seq	
pTL81	CMGC/CLK	Tb11.01.4230 Tb11.01.4250		LOF phenotype BSF by RIT-Seq	(Alsford et al., 2011)
pTL82	CMGC/CLK			LOF phenotype BSF by RIT-Seq	
pTL 7	Other/NEK	Tb927.10.5950	NEK-17	No LOF detected by RIT-Seq	(Alsford et al., 2011)
	Other/NEK	Tb927.10.5940		No LOF detected by RIT-Seq	
	Other/NEK	Tb927.10.5930		Possible LOF BSF Day 6	
pTL 4	Other/NEK	Tb927.8.1670	NEK-14	No LOF detected by RIT-Seq	(Alsford et al., 2011)
	Other/NEK	Tb927.8.1690		No LOF detected by RIT-Seq	

The expanded family of NEK kinases did not demonstrate a high number of hits during this screen. A severe growth phenotype was detected when two closely related NEK kinases, NEK12.1 and NEK 12.2 (Tb927.8.7110 and Tb924.4.5310), were jointly targeted by one RNAi fragment. Another NEK kinase, Tb927.10.14420, produced a very small reduction in growth when ablated by RNAi. The family of diverged, Orphan kinases appeared to be a poor family for providing new drug targets; only one, Tb09.211.2260, generated a loss of fitness phenotype of any note, though this was a strong phenotype and as such was investigated further.

The AGC family generated 6 hits, including PK50, a gene known to be essential in BSF parasites (Ma et al., 2010). For PK53, another essential AGC family protein kinase (Ma et al., 2010), knockdown generated an inconsistent phenotype with one cell line indicating a mild growth defect and another not. The CAMK (calmodulin -dependent protein kinase) family generated 4 hits, showing moderate growth defects. This included the lipid-droplet kinase known to generate an RNAi growth defect in BSF parasites (Flaspohler et al., 2010). The CMGC family contained a large number of protein kinases that generate severe phenotypes under RNAi induction. This included many of the CRKs that have been previously studied by RNAi. We detected CRK1 and CRK3 and CRK12 as generating an RNAi phenotype, similar to previous work (Tu and Wang, 2004; Mackey et al., 2011). The pTL analysis also detected CRK2, CRK6 and CRK 9 as having growth defects under RNAi induction. GSK3-short was detected as generating a strong phenotype (Ojo et al., 2008). Several of the MAPK family and related RCK group were identified as generating a severe growth defect (Tb927.10.5140, Tb927.10.7880, Tb927.3.690). In the smaller subfamilies of the CMGC group a dual-specificity protein kinase (Tb927.10.9600) as well as a serine-arginine rich protein kinase (Tb927.6.4970) was detected as generating intermediate growth defects.

The potential upstream activators of the MAPKs are typically found in the STE family and this provided several hits generating severe and intermediate growth defects. Tb927.10.2040, Tb11.46.0003, Tb09.211.3820, Tb927.8.5730 generated severe growth defects with Tb927.8.1100 generating an intermediate growth defect. Tb09.211.3820 and Tb927.8.5730 are STE20 family protein kinases (SLK1 and SLK2) which may represent the top of MAPK cascades (MAPK KKK). Tb11.46.0003 and Tb927.8.1100 are STE11 type protein kinases which may act as MAPKKs. Tb927.10.2040 is an undefined STE type kinase. The STE11 protein kinase Tb11.01.5650 did not generate an alamar blue value but displayed interesting morphological changes which are discussed in Chapter 5.

In the “Other” group of protein kinases the pTL screen identified two of the three Aurora kinases, Tb11.01.0330 (TbAuk1), Tb09.160.0570 (TbAuk3). CK2A1 (Tb09.211.4890) was also identified as generating an RNAi phenotype. RNAi of the predicted CAMKK (calmodulin-dependent protein kinase kinase) Tb927.10.15300 also generated a mild growth defect. The ULK homologue, and

potential regulator of autophagy Tb11.02.2050, was observed to cause a growth defect under RNAi ablation.

Several members of the pseudokinase family also showed loss of fitness phenotypes hinting at a regulatory role for these non-catalytic relatives of protein kinases. However, as they are predicted to be inactive they would make poor drug targets. The pTL screen of atypical protein kinases also yielded known essential genes such TOR1 and TOR4 (Barquilla et al., 2008, 2012).

3.3.7 Validation of mRNA Knockdown by qPCR

In order to assess the level of mRNA knockdown of the target gene a number of cell lines were assessed by quantitative PCR using the $\Delta/\Delta C_t$ method (to achieve relative quantitation). For any given target gene, primer pairs were designed for regions not targeted by the RNAi cassettes. These primer pairs were tested for efficiency using a titration of genomic DNA from 2 ng to 0.125 ng. Primers that exhibited efficiency significantly different to that of the endogenous control (which was approximately 100%) were redesigned to allow accurate comparison by the $\Delta/\Delta C_t$ method. Primers were also assessed by dissociation curve analysis to ensure only 1 PCR product was generated.

Total RNA was extracted from induced and uninduced cultures, treated with DNaseI to remove genomic DNA and then converted to cDNA using reverse transcriptase with random hexamer primers. Control reactions lacking reverse transcriptase were included to allow detection of gDNA contamination in the samples. cDNA reactions were used as templates in qPCR reactions. GPI8 primers were used in endogenous control reactions and relative mRNA values were determined using the $\Delta/\Delta C_t$ method using the uninduced, reverse transcribed templates as the reference sample.

Transcripts were significantly reduced for 8 of the 9 genes studied by qPCR (Figure 3-7). mRNAs for target genes were typically reduced by 40-60%. RNAi of CRK3 using the original pRPA^{ISL}/pGL1875 and the pTL-based vector were not significantly different. RNAi of MKK5 did not lead to a significant reduction in mRNA levels, though a downward trend was observed. Further validation of knockdown for a NEK12.1, ZFO, and DFK genes was carried out using western

blotting to assess protein levels of endogenously tagged alleles (see following chapters).

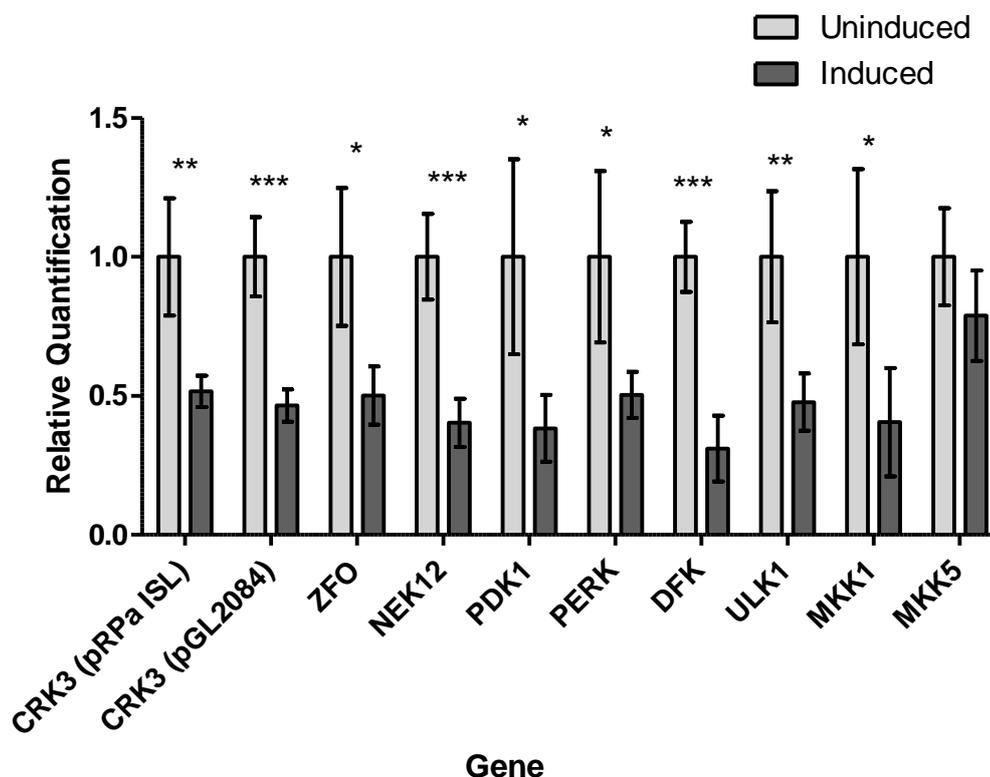


Figure 3-7: Bar chart showing the mean relative quantification of mRNA in induced and uninduced RNAi cell lines. Error bars represent the standard deviation of RQ values from the 4 technical replicates used to generate the mean. Significance of the result (as derived from unpaired t-test) is indicated above bars for each gene (*: p 0.01 to 0.05, **: p0.001 to 0.01, *** : p< 0.001). Identifiers for genes can be found in Table 3-1, Table 3-2 and Table 3-3.

3.4 Discussion

3.4.1 Comparison of pTL Screen to Candidate Studies

The pTL RNAi libraries efficacy was validated by its ability to rediscover previously confirmed phenotypes (Table 3-5). The initial validation of the system was performed against CRK3, a known essential gene in trypanosomes (Tu and Wang, 2004). Other candidate genes such as TLK1 and CK1.2 were detected as essential (though the pTL RNAi also co-targeted their related genes TLK2 and CK1.1 respectively.) Detailed studies have determined the importance of TLK1 and CK1.2 individually, so pTL RNAi constructs targeting them individually were

not developed (Li et al., 2007a; Urbaniak, 2009). Several genes which have been knocked out of monomorphic BSF parasites did not generate growth defects when targeted in RNAi lines; these include ZFO, MKK1, MKK5, MAPK 2, MAPK 4 and MAPK 5 (although for MKK5 mRNA levels were not detected as significantly reduced by qPCR). This confirmed the prediction that double stranded RNAi coming from the transcribed attL sites had no disruptive effect on trypanosome biology. These data combined with the qPCR validation suggest that the pTL RNAi is operating effectively and reliably, allowing confidence to be placed in the accuracy of this screen.

Table 3-5: Table of pTL phenotypes confirming previously published work on BSF trypanosomes.

Kinase	pTL Phenotype	Published BSF Phenotype	Ref.
CRK3		Essential by RNAi	(Tu and Wang, 2004)
Auk1		Essential by RNAi	(Li and Wang, 2006)
TLK1		Essential by RNAi	(Li et al., 2007b)
CK1.2		Essential by RNAi	(Urbaniak, 2009)
GSK3-s		Essential by RNAi	(Ojo et al., 2008)
PK50		Essential by RNAi	(Ma et al., 2010)
PK53		Essential by RNAi	(Ma et al., 2010)
LDK		Mild growth defect RNAi	(Flaspohler et al., 2010)
TOR1		Essential by RNAi	(Barquilla et al., 2008)
MKK1		Non-essential by double knockout	(Jensen et al., 2011)
MKK5		Non-essential by double knockout	(Jensen et al., 2011)
MAPK 2		Non-essential by double knockout	(Müller et al., 2002b)
MAPK 4		Non-essential by double knockout	(Güttinger et al., 2007)
MAPK 5		Non-essential by double knockout	(Domenicali Pfister et al., 2006a)
ZFK		Non-essential by double knockout	(Vassella et al., 2001)

It should be noted that in some instances (notably for the several metacaspases) performing a gene knockout can show a gene is not essential whereas RNAi can generate a strong phenotype; however, this was not observed in this study for protein kinases that have been knocked out in previous works (Helms et al., 2006; Proto et al., 2011). Also the strain of trypanosomes can have an effect on

the observed phenotype. Pleiomorphic trypanosome lines (capable of efficient tsetse fly transmission) can show deleterious effects of genetic interference whilst monomorphic lines do not (Vassella *et al.*, 2001). Deletion of ZFK causes no defect in the proliferation of monomorphic MITat 1.4 cell line but leads to a growth defect in AnTat 1.1 after about 4 days in culture. The 2T1 strain is considered monomorphic and may give an underrepresentation of genes that may be essential to pleiomorphic cell lines.

The study by Mackey *et al.* generated an RNAi library targeting 30 genes in the BSF trypanosome kinome and analysed the cell lines using a luciferase based HTS assay (Mackey *et al.*, 2011). However, the results correlated poorly with the pTL Screen data and the RIT-seq data. mRNA downregulation was only confirmed for 2 genes and the study was conducted using the pZJM system which is now considered a less reliable method for performing RNAi in BSF trypanosomes due to the range of sites it can integrate into the genome (Mackey *et al.*, 2011).

3.4.2 Comparison to RIT-Seq Data

The genome-wide study for loss of fitness phenotypes by Alford *et al.* has provided a large data set against which to compare our kinome-wide RNAi. For the majority of the genes studied the same phenotype was found. To simplify the analysis, RNAi targets against multiple genes were ignored. For the remaining 177 genes agreement on phenotype was found for 129. RIT-Seq did not identify a phenotype in a group of 27 genes for which the pTL screen generated a loss of fitness phenotype. This may be due to the RIT-seq approach lacking the resolution to detect some RNAi phenotypes if that region of the genome is under-represented in the RNAi library. Differences in the strain of parasite used and the RNAi system (with the stem-loop being more effective in the pTL screen) may account for the observed differences.

Twelve genes identified in the RIT-Seq data as generating loss of fitness phenotypes at Day 6 were not identified in the pTL screen. The pTL screen analysed cell lines at 72 hours, as we were aiming to detect strong phenotypes that may be more attractive drug targets. Thus we may have not detected these more subtle phenotypes that only become apparent after 72 hour. RNAi phenotypes requiring longer to appear may indicate that the kinase in question

is not an ideal drug target. The alamar blue assay required cells to be seeded at 2×10^4 cells per ml and for 72 h- this may have masked very minor phenotypes. Performing growth curves of these cell lines only identified one that generated a growth defect after 6 days, Tb927.2.1820 (M Saldivia, Unpublished).

Possibly the most relevant cohort for scrutiny is the 9 genes for which RIT-seq generates a strong phenotype but the pTL screen does not. These may represent false negatives in the pTL screen or false positives in the RIT-seq screen. One example is the well characterised gene *PLK* (see below).

3.4.3 False Negatives

As with any large-scale screen, we expected some false negatives. This was clearly demonstrated by our inability to generate an RNAi phenotype for *PLK*, where RNAi should lead to cytokinetic defects and growth arrest within 8 hours (Hammarton et al., 2007). This was despite targeting the gene with two separate RNAi fragments in multiple different transfections. Initially it was assumed that the RNAi could be “leaky” while selecting for transfectants, however, we could still generate RNAi lines for other critically essential genes. The same batch of serum was used for all transfections (Gibco 10270), is marketed as tetracycline free and transfections for other essential genes proceeded unhindered; so this was not believed to be causing expression of the RNAi cassette. Quantification of *PLK* mRNA knockdown has not yet been performed and as such there is no data to show whether the RNAi is actually downregulating the gene.

A small cohort of 9 genes were detected by Alford *et al.* as having a loss of fitness phenotype after 3 day RNAi in BSF parasites and were not detected in the pTL screen. These genes have not been independently validated but the RIT-seq data suggest that these may be false negatives in our screen. However, these could also be false positives in the RIT-seq data due to low sequence coverage over that particular gene; for example, cell density at time of induction may also play a role in the detected phenotype. TOR 2 RNAi did not produce a growth defect by alamar blue assay or RIT-seq but when growth curves were performed after seeding the cells at 10^4 per ml a growth defect replicating the published phenotype was observed (W. Proto, Unpublished) (Barquilla et al., 2008).

The 9 genes in question have not been investigated further; however, it would be prudent to check whether mRNA is being degraded in these cell lines to begin to establish why the RNAi is not effective.

3.4.4 Other high-throughput cloning systems

Recently, during the course of this research, other systems have been generated to increase the efficiency of cloning stem-loop RNAi vectors. A Gateway® system named pTrypRNAiGate has been utilised to facilitate this (Kalidas et al., 2011). The Gateway system from Invitrogen uses an Integrase from λ phage to catalyse the recombination of specific DNA sequences termed att sites. It is not dependent on restriction enzymes sites and uses counter selectable markers to ensure very high cloning efficiencies (Fu et al., 2008). However pTrypRNAiGate first requires the cloning of a PCR product into a donor vector and then a second recombination reaction to clone the insert into the RNAi expression vector and does not bring any reduction in the number of sub-cloning steps over the restriction enzyme method. This vector randomly integrates into the rRNA spacer regions of the parasite and does nothing to resolve the problem of inter-clonal variation. Another recent method utilised complementary restriction sites and oligos that include random sequence to introduce the loop into the stem-loop construct (Atayde et al., 2012). This approach is a valid one for all stem-loop RNAi constructs and requires only one PCR product to be generated and one cloning step to take place but still relies on restriction enzymes and is affected by their associated limitations such as reduced cloning efficiency.

3.4.5 Further Utilities of the RNAi Library

The pTL library has been used for a relatively simple screen, generating information only for cell viability phenotypes. This is due to the recurring problem that phenotyping large numbers of cell lines by multiple techniques is very labour intensive. New technologies may allow an increase in the throughput of phenotypic analysis of libraries like this. High-content imaging systems should allow multiple phenotypic characterisation of large number of cell lines subjected to multiple variables. This is termed High-Content Screening (HCS). At a simple level, HCS can be used to look for changes in fluorescence whether this is from a probe or from a fluorescent reporter gene. Increasing the complexity of

analysis can allow the counting of organelles within a cell, the infection efficiency and intensity of an intracellular parasite or even morphological changes of cells and organelles. Using computational machine learning or writing scoring algorithms is a crucial part of the analysis. The ability to prepare samples consistently, on a large scale, and to image the results of an assay require appropriate robotic systems in aseptic conditions (Conrad and Gerlich, 2010). The installation of the Scottish Bioscreening Facility (SBF) at the University of Glasgow provides all of this and presents an ideal opportunity for implementing the pTL library into HCS assays. The SBF contains a GE Healthcare IN Cell 2000 automated microscope system, Beckman BioMek liquid handling robotics and various peripherals for dispensing medium, cell cultures and compound libraries. An example is presented in Chapter 5 about how this can system can be used for dissecting BSF to PCF differentiation of African trypanosomes.

Trypanosomes have a carefully choreographed cell division cycle, and multiple protein kinases have already been implicated in regulating it in both PCF and BSF parasites. Screening the pTL library for cell cycle defects could uncover novel regulators of the BSF cell division cycle that are not conserved in other organisms. RNAi could be induced in the pTL cell lines and these then subjected to cell cycle analysis by DAPI counting. This is time consuming, tedious and has potential for investigator bias to creep in; thus it is an ideal target for automation. It has already been automated using conventional microscopy and ImageJ scripting (Wheeler et al., 2012). Thus, it is not unreasonable to imagine this being translated into a HCS workflow allowing large-scale cell cycle analysis to be possible.

HCS assays can be performed for any phenotype that can be assayed using a fluorescent reporter. The CRK12 kinase has been shown to play some role in endocytosis as when ablated by RNAi cells exhibit accumulation of material in the flagellar pocket, showing a “big-eye” phenotype (C. Costa, pers. comm.). This is assessed by using fluorescently labelled ConA or FM4-64 followed by microscopy (Field et al., 2004). Other RNAi lines could be screened for this defect using HCS microscopy to identify other kinases acting on this process, and that could be part of a pathway regulating or being regulated by CRK12.

Treating cells with fluorescent probes (for example a fluorescently-labelled antibody) can add complexity to the sample preparation in a HCS. Reporter cell lines can relieve this to some extent, by generating a fluorescent readout from reporter genes or fluorescently labelled fusion proteins. One investigation, which is being used in complementation with the pTL library, is looking into the process of autophagy in trypanosomes. Autophagy is a process undertaken by cells, often during differentiation or when under nutrient stress, where a double membrane vesicle is constructed around a substrate or organelle that is then targeted for degradation in the lysosome (Brennand et al., 2011; Duszenko et al., 2011). A marker for mature autophagosomes is the ATG8 protein, of which *T. brucei* contains two forms. Fusion proteins of YFP-ATG8.1 and YFP-ATG8.2 have been created and expressed in BSF form trypanosomes. The aim is now to screen the trypanosome kinome for protein kinases that can induce or suppress autophagy. The ATG8 reporter genes are compatible with the pTL system and have been transfected into several cell lines for validation. Normal BSF cells very rarely possess autophagosomes but when TOR1 is depleted by RNAi the number increases, as has been shown by electron microscopy. In a cell line containing the ATG8.2 reporter cell line and the pTL149 which targets TOR1, the number of autophagosomes was observed to increase (W. Proto, Unpublished). This appears to validate the work by Barquilla *et al.* (2008) and also the feasibility of using the pTL system along with this reporter, although large numbers of transfections would be required to combine the reporter plasmid with the pTL library. It does appear that this methodology would be compatible with a HCS study.

More speculatively, the system could be used alongside protein kinase inhibitors to look for synthetic-lethal phenotypes. The concept of an essential gene is important. Many genes are not individually essential, in brewer's yeast for example, in a rich growth medium only 17% of genes were essential for viability, though cumulatively, 40% did give a detectable loss of fitness in either rich or minimal media (Giaever et al., 2002). In the RIT-seq analysis of *T. brucei* only 10% of genes were essential in all conditions tested, and 43% of knockdowns did not give a loss of fitness under any conditions tested (Alsford et al., 2011). In the analysis of Chromosome 1 by RNAi only 23% of genes gave growth defects (Subramaniam et al., 2006). The pTL analysis gives an essential or important role for 27% of the kinome. Aside from any experimental problems the reason for

some of the “non-essential” genes being non-essential is potentially due to either direct redundancy of the gene product or the ability of other biological pathways to accommodate its absence. Cell signalling is probably not best viewed as a linear pathway, but a complex network. Removal of one of the nodes in a pathway or network may shift signalling in an alternative route to ensure the survival of the organism (D’Elia et al., 2009; Brognard and Hunter, 2011). A potential way of probing these networks is through synthetic-lethal studies. By perturbing two genes which are not individually essential a lethal phenotype can be achieved due to an inability to compensate for the double perturbation. By taking the STL lines for which RNAi did not generate a growth defect and then inhibiting another gene either by double RNAi or by adding inhibitors one can begin to probe for association between genes that are not individually essential. Another strategy could be to knock a non-essential gene of interest out of the 2T1 parental cell then transfect the pTL plasmids for non-essential protein kinases and screen for defects (sufficient drug selectable markers do exist to do this). In doing so, one can begin to identify signalling networks essential to an organism or disease. The concept of synthetic lethality is being increasingly investigated as a possibility of designing effective combination therapies to improve the treatment of various cancers.

The use of the pTL system is ideal for exploring the biology of BSF trypanosomes but due to the characteristics of the 2T1 cell line the study of differentiation and procyclic trypanosomes is complicated. The 2T1 cell line will not differentiate into a viable PCF line and will not transmit through tsetse flies. The library is therefore unsuitable for studying the role of protein kinases in the progress of the parasite through the tsetse fly. Although not relevant for drug target validation, a study of the role of protein kinases may assign a role to some of the protein kinases that were shown not to be essential in BSF. This type of systematic study of the role of kinases in life cycle progression has been carried out in the rodent parasite *Plasmodium berghei* (Tewari et al., 2010). Modifying a pleiomorphic strain (such as Antat 1.1) to be compatible with the pTL RNAi library might allow such a study.

In summary the pTL RNAi screen has genetically validated a cohort of protein kinases as potential drug targets. Further validation of each of these is necessary to classify them as a true drug target, for example testing the essentiality of

kinase activity of each enzyme. The library provides a resource for future studies into the biology of *Trypanosoma brucei* and could allow chemical genetic screens or HCS studies of various cellular processes.

4 *Trypanosoma brucei* Protein Kinases as Drug Targets

4.1 Targets that were identified in the screen for further study

The trypanosome kinome possesses several families of protein kinases that are considerably different to those found in the human kinome and therefore may present opportunities for selective inhibition. The family of NEK protein kinases, which contains protein kinases resembling those found in humans, is dramatically expanded in *T. brucei*, potentially hinting at their importance. The other family of trypanosome protein kinase that are interesting from a drug target validation viewpoint are termed Orphan protein kinases, as they have very divergent amino acid sequences such that no orthologues have been identified in protein kinase families in other organisms. Trypanosomes also possess protein kinases with homology to human protein kinases, such as PDK1, that are the focus of drug development programmes - this may present opportunities for repurposing existing inhibitors. This chapter discusses the validation of NEK, Orphan and PDK1-related protein kinases as drug targets. The main focus of the work is a NEK kinase predicted to have an active site amenable to rational drug design.

4.2 Introduction

4.2.1 NEK Family Protein Kinases

NEK stands for Never-in-Mitosis Related Protein Kinase. NEK kinases are serine/threonine kinases and *Trypanosoma brucei* is predicted to possess an expanded family of this type of protein kinase (Previous Page - Figure 1-7). The first discovered NEK protein kinase was NimA, for which the family is named (Osmani et al., 1988). It was discovered during a screen for mitotic defective mutants of the filamentous fungus, *Aspergillus nidulans*. Mutants were assessed for being blocked-in-mitosis (bim) or never-in-mitosis (nim). The gene from the first nim mutant, *nimA*, was investigated using a regulatable system for further analysis. Conditional mutants of the *nimA* gene could be blocked in the G2 phase of the cell cycle at restrictive temperatures. The role of this gene in regulating mitosis was further confirmed by overexpression of the gene. Cloning of the

cDNA followed by sequencing showed that the protein product possessed the characteristic of a protein kinase (Osmani et al., 1988). Subsequently homologues were identified in other organisms (hence the use of NEK nomenclature), with further characterisations showing NEKs have a diverse range of cellular functions. NEKs are a deeply conserved family of protein kinases with phylogenetic inferences suggesting that the last common ancestor of eukaryotes (which had to be a dividing, ciliated cell) had at least 5 NEKs (Parker et al., 2007). As some of the mammalian NEKs have been implicated in cancer and other diseases their biological roles have been studied, giving insights into how they function. The literature on this subject is expanding rapidly and is reviewed well in (Quarmby and Mahjoub, 2005; O'regan et al., 2007; Moniz et al., 2011).

The NEKs that have been currently characterised appear to fall into two functional clades; those that regulate mitosis and those that regulate cilium development. The *nimA* homologues were the first to be studied in other organisms and Nek2 was the first to be researched in humans. Direct functional homology amongst the *nimA* sequence homologues from different organisms appears to be rare. *nimA* deficiency in *Aspergillus* cannot be rescued by *fin1*, KIN3 or NEK2 (the *nimA* sequence homologues in *S. pombe*, *S. cerevisiae* or mammals respectively) but only by the *nim-1* from *Neurospora crassa*, a related fungus (Pu et al., 1995). Following *nimA* characterisation the main focus of research on NEKs was conducted on HsNEK2 due to its roles in mitosis. During mitosis HsNEK2 regulates centrosome segregation. It localises to centrosomes and phosphorylates centrosomal proteins such as cNAP-1, Rootletin and β -catenin. Inhibition of HsNEK2 prevents centrosome segregation, spindle assembly and formation of multinucleate cells. It is highly essential for mouse embryogenesis. Increased levels and activity of HsNEK2 are detected in a large number of cancer types and cancer cell lines. HsNEK2 siRNA silencing reduces the proliferation of a number of these cell types suggesting HsNEK2 presents a drug target for anti-cancer therapy (Kokuryo et al., 2007; Moniz et al., 2011).

In humans three other NEKs have been shown to participate in regulating mitosis; interestingly, these operate in a cascade. HsNEK6 and HsNEK7 are 87% identical within their kinase domains and do not possess much sequence similarity outside this domain. These appear to regulate the mitotic spindle and

centrosome splitting and cytokinesis. They are activated by NEK9, which in turn is activated by CDK1 and PLK1. The activity of HsNEK7 and HsNEK7 on the kinesin Eg5 causes it to accumulate at the centrosomes prior to centrosome separation. These are microtubule dependent processes and it appears that several NEKs can regulate microtubule dependent events outside of mitosis (Moniz et al., 2011). HsNEK3, for example, plays a role in regulating acetylation of microtubules in neurons. When expression of an inactive mutant is induced in cultured human neurons the cells lose their polarity and have abnormal morphology (Chang et al., 2009).

The other functional clade of NEKs is made up of those that regulate the development of cilia and flagella. This was first highlighted by studies implicating NEK1 and NEK8 (in mice and zebrafish) disruption in the deregulation of primary cilium. This is a sensory organelle and when disrupted by ablation of NEK1 and/or NEK8 leads to polycystic kidney disease (Mahjoub et al., 2005). Interestingly, NEKs seem to be expanded in ciliated and flagellated organisms and their importance to these organisms may be related to regulating the formation of these organelles. It is likely that NEKs evolved and diverged early in the evolution of eukaryotes (Parker et al., 2007). *Chlamydomonas reinhardtii*, a bi-flagellated, single-celled algae, possesses 11 NEKs. Two NEKs, Fa2p and Cnk2p, localise to the flagellum are responsible for regulating the length of its flagellum and the size of its cells and localise to the flagellum (Mahjoub et al., 2004; Bradley and Quarmby, 2005). The family of NEKs is also expanded further in the freshwater ciliate *Tetrahymena thermophila*. This organism is a single cell possessing several hundred cilia and it has a complement of 39 NEK kinases. A number of these have been localised to distinct subsets of cilia and three have been implicated in several different mechanisms of cilia disassembly (Wloga et al., 2006).

In an extreme example the NEK family of protein kinases is expanded in a most striking manner in *Giardia*. The intestinal parasite *Giardia lamblia* has a huge expansion and diversity in its NEKs. In the WB strain 198 proteins have been classified as NEKs, though, only 56 of these are predicted to be active. The 198 proteins make up 71% of its kinome and 3.7% of its proteome. When these expanded families of NEK kinases were compared between 3 strains of *G. lamblia* it was apparent that a core of genes were conserved but many were

strain specific loss or gains and many were diverged. It is suggested that these differences may account for strain specific differences (Manning et al., 2011). Characterisation of some *Giardia* NEKs has been undertaken. Homologues of HsNEK1 and HsNEK2 genes were identified in *Giardia* by sequence similarity and termed GINEK1 and GINEK2 respectively. These were shown to be active kinases *in vitro*, with activity against Histone H1 and Histone H3, and Histone H1 and BSA respectively. As targeted gene knockout or inducible RNAi silencing is not feasible *Giardia*, the two NEK kinases were overexpressed in kinase-dead form in order to achieve a dominant negative phenotype. Both led to growth defects but it demonstrated that GINEK1 was involved in regulating cytokinesis, as there was a large increase in cells stalled in cytokinesis in the cultures. When localised by immunofluorescence microscopy it was shown that the two protein kinases localised to areas of intense microtubule remodelling such as the mitotic spindle poles and microtubules, and the microtubules of the ventral disc. They also show some co-localisation with phosphorylated aurora kinase. Overexpression of GINEK1 and kinase dead GINEK2 led to a drop in the number of excyzoites that could be obtained from cysts formed by the relevant overexpression cell line, showing these kinases play some role in excystation, a process that requires dramatic remodelling of the cytoskeleton as well as segregation of nuclei into daughter cells (Smith et al., 2012).

In comparison with flagellated parasites only four NEK kinases have been identified in the apicomplexan parasites *Plasmodium falciparum* and *P. berghei*, which have been characterised in some detail. Pfnek-1 is the only one expressed in asexual parasites and is an active kinase *in vitro*. It can phosphorylate Pfmap2 (MAP kinase 2) *in vitro* leading to the activation of this kinase (Dorin et al., 2001). It is essential to the multiplicative erythrocytic stage and is expressed in male gametocytes, making it a therapeutic and transmission blocking target (Dorin-Semblat et al., 2011). Pfnek-3 is a dual specificity Ser/Thr and Tyr kinase has also been shown to phosphorylate and activate Pfmap2 in its activation loop *in vitro* (Lye et al., 2006; Low et al., 2007, 2012). In *P. falciparum* *nek-2* and *nek-4* are individually essential for DNA replication prior to meiosis in the zygote (Reininger et al., 2009; Tewari et al., 2010). The NEKs of *P. falciparum* have been shown to possess specificities for different generic substrates *in vitro* kinase assays.

Trypanosoma brucei has 21 genes predicated to encode NEK kinases compared with 15 in humans. NEKs from *T. brucei* do not form close relationships to the NEKs from *P. falciparum* nor to *S. cerevisiae*, *D. melanogaster* or *C. elegans* in phylogenetic analyses (Parsons et al., 2005a). However, some of the NEK kinases in the latter organisms do form close relationships, suggesting divergence of trypanosomatid NEK kinases from other eukaryotes. Most of the of *T. brucei* NEKs possess the architecture of human NEKs, with a protein kinase domain located at the N-terminus or close to it and a C-terminal extension containing other domains. Several contain coiled coil domains C-terminal to the active site, reminiscent of HsNEK1 an HsNEK2 (Figure 4-1).

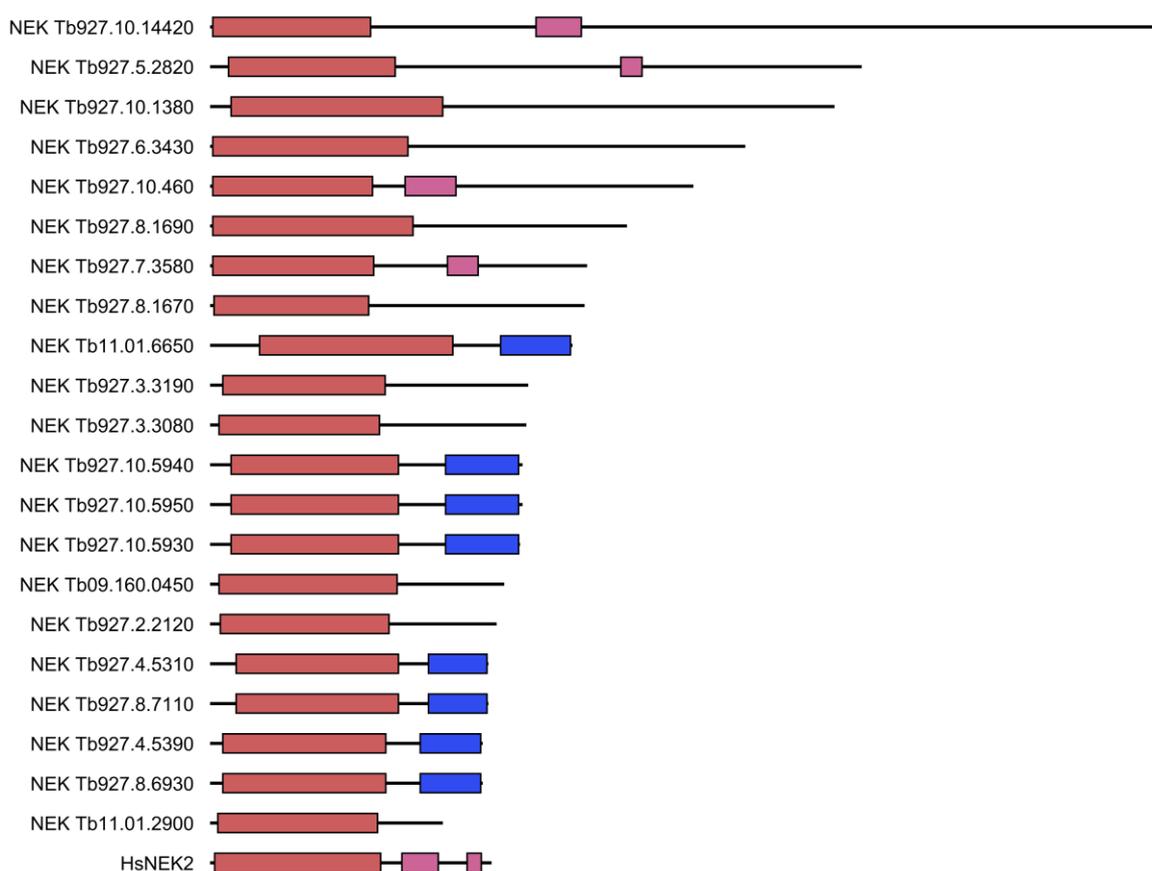


Figure 4-1: Depiction of the domain architecture of the predicted NEKs of *T. brucei*. Protein kinase domains are shaded in red, coiled coil regions in pink and PH domains in blue.

The NEKs of *T. brucei* have not been investigated in much detail at all, only several have had any characterisation and only one investigated in detail (in procyclic forms). The first to be identified were NRKA and NRKB, which were cloned using degenerate primers for protein kinase domains, in the pre-genomic era (Gale and Parsons, 1993). Analysis of the sequence showed that outside the

catalytic domains the protein showed very little homology to HsNEK1 and NIMA, so the researchers concluded that NRKA and NRKB were not the functional homologues of the former; however, no functional characterisation was undertaken. The transcription of the genes were analysed by northern blot and it was shown to be constitutively expressed in both strains at both stages of the life cycle (Gale and Parsons, 1993). The genes were later shown to be translationally regulated, even though the mRNAs are constitutively expressed. The protein levels and kinase activity were increased in progression from slender BSF to stumpy BSF and the highest level of expression took place at this point. No work was done to identify the location at which they act in the cell.

NRKC has had much more characterisation. Pradel *et al.* (2006) demonstrated that NRKC has a high similarity to human NEK 2. It is expressed as a protein in procyclic and bloodstream forms. *In vitro* kinase activity was demonstrated from recombinant protein, with the generic substrates α/β -casein being preferentially phosphorylated. Using fluorescent antibodies and C-terminal GFP-conjugate TbNRKC demonstrated that it localises to the basal body in PCF forms (a microtubule organising centre), and with tyrosinated tubulin. This suggests that TbNRKC could have some role in organising microtubules. RNAi of TbNRKC in PCF led to an accumulation of cells with 4 basal bodies but was not lethal nor mildly deleterious. The overexpression of a kinase-dead mutant protein also led to the accumulation of a 4-basal body cell type. However, when active TbNRKC was overexpressed this led to a lethal cessation of cell division. Cells showed clusters of numerous, non-flagellated basal bodies. This blocked kinetoplast division and cytokinesis. The researchers postulate that TbNRKC is responsible for activating the dissociation of links between dividing basal bodies (Pradel *et al.*, 2006).

4.2.2 PDK1

In mammals, PDK1, or 3-phosphoinositide-dependent protein kinase-1, is member of the AGC family of protein kinases. It is a constitutively active protein kinase that regulates the activity of a number of other AGC kinase, and as such is termed a “master kinase”. It functions in the transduction of extracellular stimuli to intracellular signalling pathways. Following stimulation of cells with insulin or growth-factor, PI3 kinases (phosphatidylinositol 3-kinases) become activated and generate phosphatidylinositol 3,4,5-trisphosphate in the cell

membrane. Both PDK1 and PKB can bind phosphoinositides, using C- and N-terminal PH domains respectively, and this co-localisation allows PDK1 to phosphorylate PKB in its activation loop. Other AGC kinases such as PKA, SGK and S6K, possess a C-terminal hydrophobic motif that PDK1 can bind to, allowing it to activate them by phosphorylation outside of the activation loop. The region in PDK1 responsible for this binding is termed the “PIF-Pocket” (Mora et al., 2004; Pearce et al., 2010).

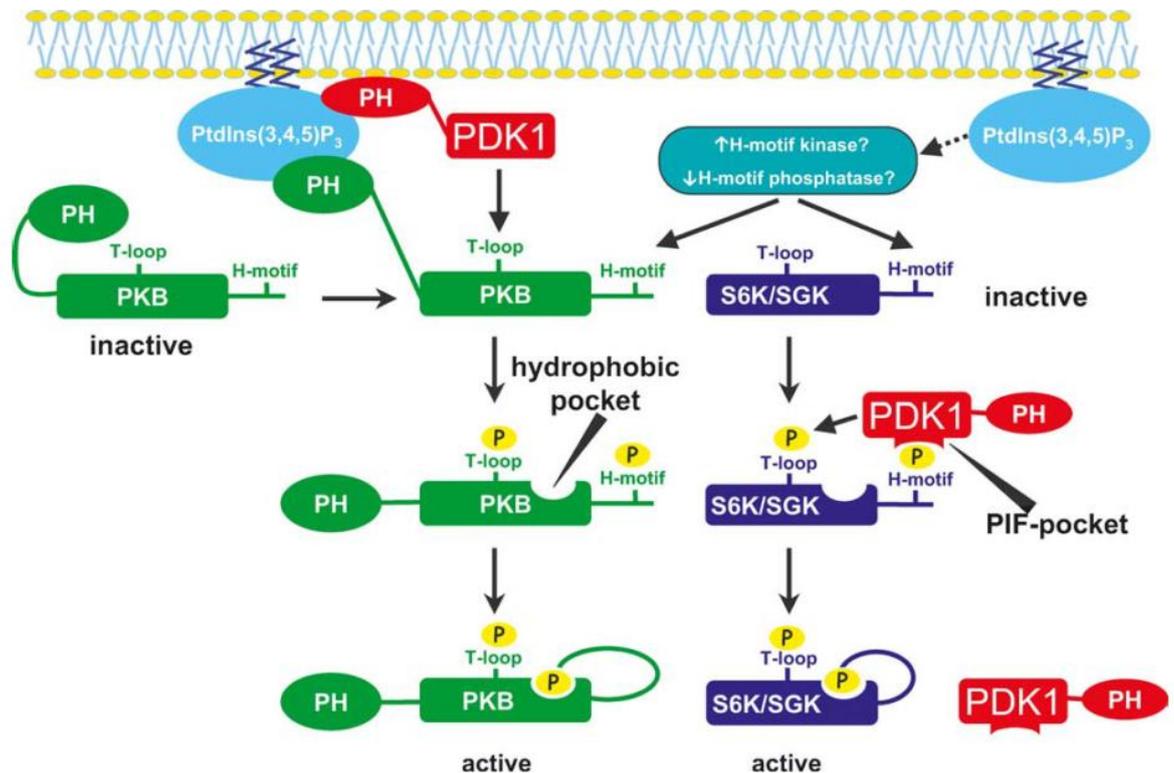


Figure 4-2: The mechanisms of PDK1 activation of PKB, SGK and S6K. Both PKB and PDK1 are localised to the plasma membrane by binding to phosphoinositides using their PH domains. PDK1 phosphorylates the activation loop of PKB allowing a hydrophobic pocket to accommodate the hydrophobic motif (phosphorylated by another protein kinase) and locking PKB in an active conformation. PDK1 can bind SGK and S6K by accommodating their phosphorylated hydrophobic motifs into its PIF pocket thus allowing it to phosphorylate the activation loop of these two protein kinases. From (Mora et al., 2004).

Increased levels of phosphoinositides in certain cancer cells can drive their proliferation through the upregulation of PKB and SGK activity. As PDK1 integrates this signalling it is being investigated as a target for therapeutic inhibition, with a large number of inhibitors being tested against PDK1 (Peifer and Alessi, 2008; Raimondi and Falasca, 2011). Due to the potential of there

being inhibitors of TbPDK1 in the vast number of compounds developed against HsPDK1 this protein kinase was investigated as a target that would be possible for inhibitor repurposing.

4.2.3 Orphan Kinases

The orphan kinases of *T. brucei* show no clear homology to any currently identified families of protein kinase, despite being readily identifiable as protein kinases. They were the most divergent families between the three trypanosomatid kinomes defined by Parsons *et al.* and show poor conservation between the three species studied, representing lineage specific gains and losses (Parsons *et al.*, 2005a). Due to their increased sequence divergence from any protein kinases found in humans they may possess structural differences in the active site that can be exploited in the design of selective inhibitors. This potentially increases the chemical space within which selective inhibitors could be developed for any essential orphan kinases making them possible good drug targets.

4.2.4 Research Aims

The objectives of the work presented here was to further investigate the validity of three protein kinases as potential drug targets. This included genetic validation of the protein kinases by RNA interference, generation of recombinant protein for chemical validation and molecular modelling to gain information on their structure. A protein kinase from the expanded NEK family was analysed in detail along with a diverged Orphan protein kinase and a potential homologue of PDK1.

4.3 Results

4.3.1 Features of NEK12

After the initial definition of the *Trypanosoma brucei* kinome, it was possible to detect by multiple sequence alignment that one NEK kinase (Tb927.8.7110) contained a small gatekeeper residue (J. Mottram, Unpublished). Tb927.8.7110 (NEK 12.1) has a paralogous gene Tb927.4.5310 (NEK12.2) that is the result of a chromosome translocation. The coding sequence of the 1323 nt genes share 87%

identity though the UTRs are more diverged at 36% and 24% for the 5' UTR and 3' UTR respectively (Jackson, 2007). The products of the two genes are also highly similar, at 85% identity of the amino acid sequence (Figure 4-4). Both proteins are 440 amino acids in length, NEK12.1 is predicted to have a mass of 49.6 kDa and NEK 12.2 a mass of 50.2 kDa. The proteins have an N-terminal kinase domain followed 48 amino acids later by a C-terminal PH domain. The gatekeeper residues of the two active sites are Ala117 in NEK 12.1 and Met117 in NEK 12.2; alanine possesses the second smallest side chain of the amino acids (after glycine). It was predicted that this would generate a large active site in NEK 12.1 compared to NEK12.2 and other kinases in *Trypanosoma brucei*. This was further investigated by molecular modelling. Three phosphopeptides corresponding to two phosphosites (Ser 195 and Ser 197) in the activation loop of NEK12 were detected by Nett *et al.* 2009. This was assigned to both kinases in their analysis but this should not be assigned to one or the other as the sequence the peptide covered was identical in each kinase (Nett *et al.*, 2009b). These key features are annotated to a schematic representation of NEK 12.1 in Figure 4-3. The catalytic lysine was identified as K70, and was considered as a site for mutagenesis to create kinase-dead mutants.

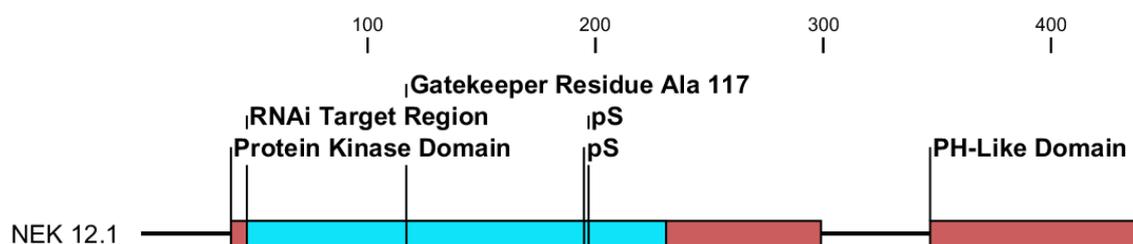


Figure 4-3: Schematic representation of NEK 12.1 protein domain architecture with key features annotated. This representation is the same for NEK 12.2 with the exception of Met 117 replacing Ala 117. The scale above the diagram shows the amino acid number. The equivalent region of the CDS targeted by RNAi is highlighted in cyan. Ser 195 and Ser 197 have been detected as phosphorylated by Nett *et al.* 2009.

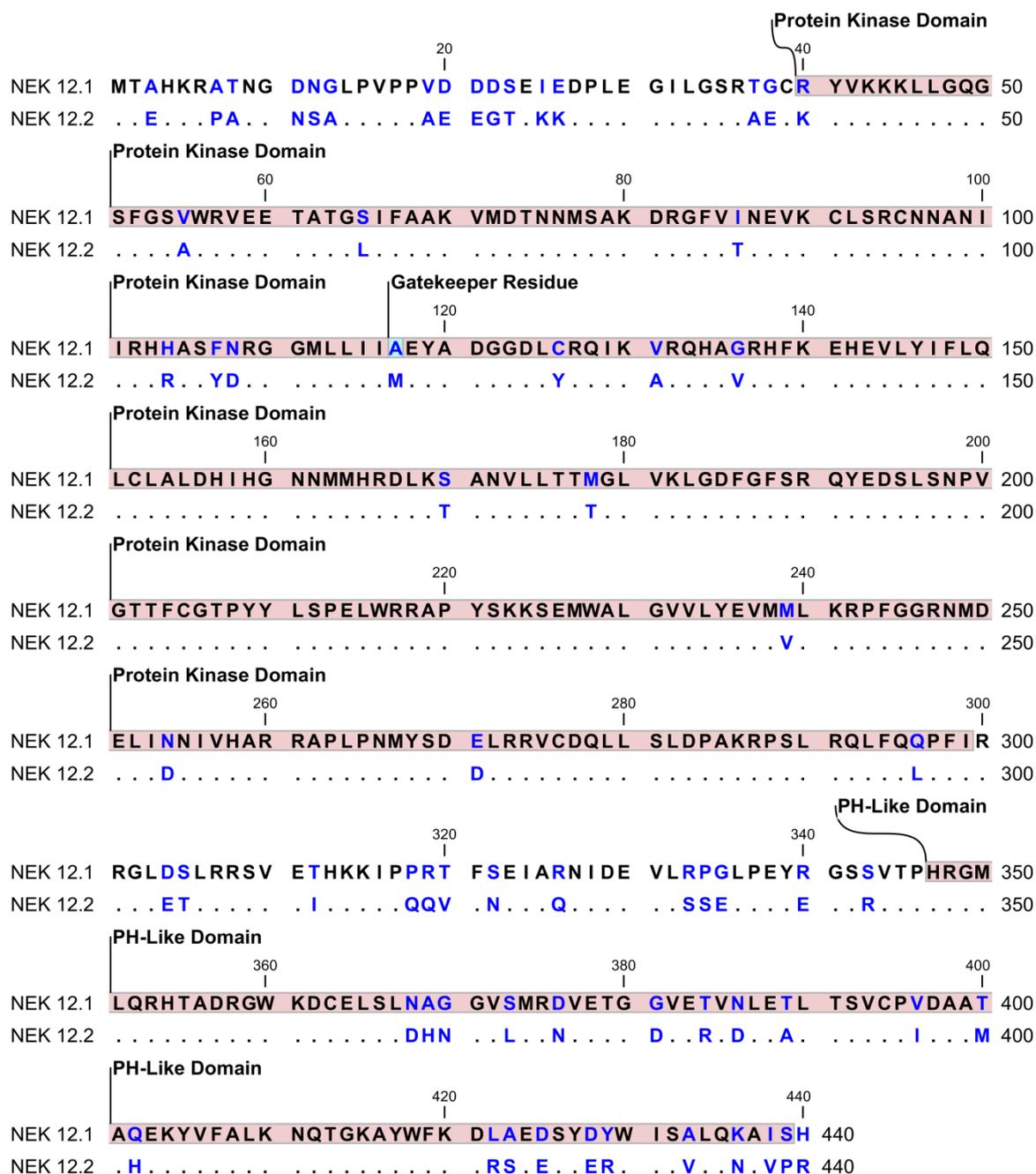


Figure 4-4: Sequence alignment of NEK 12.1 and NEK 12.2 proteins. The protein kinase and PH domain are highlighted by red bars. The position of the gatekeeper residue is annotated and differences in amino acid are highlighted in blue, identical residues are indicated by dots.

4.3.2 NEK 12 RNAi

To assess the relevance of the NEK 12 protein kinases a stem loop RNAi construct (pGL2064) was generated and transfected into *T. brucei* 2T1 BSF cells. Two clones (15356/15357) were selected for investigation. The RNAi plasmid was generated targeting 555 nt of the coding sequence, which results in cross-targeting of the two NEK 12 mRNAs leading to a double knockdown. Due to the

similarity of the two NEK 12 coding sequences this was unavoidable as the UTRs were also unsuitable for RNAi targeting due to their similarity to other UTRs in the *T. brucei* genome.

Initial assessment of the NEK12 RNAi phenotype was carried out by generating a cumulative growth curve of the BSF RNAi phenotype *in vitro*. Ablation of both NEK12 proteins by RNAi generated a severe growth defect after 24 h. After 24 h there was a rapid cell death phenotype in the induced cultures and by 96 h no viable cells were visible in the medium. This was observed in two biological replicates (and three technical replicates) (Figure 4-5a).

In order to assess the severity of NEK 12 ablation in a more realistic situation *in vivo* RNAi studies were carried out in ICR mice. The 2T1 strain of *T. brucei* generates acute infection dynamics due to lack of density regulation by the parasites. Mice were inoculated on Day 0 with 5×10^4 parasites (clone 15356) and 24 hours later the test mice were presented with doxycycline laced sugar water. Parasitaemia was monitored every 24 h by haemocytometry on blood obtained from tail venepuncture. Mice were typically culled if parasitaemia rose above 10^9 parasites per ml of blood or after 96 h since inoculation. The parasitaemia of control and RNAi induced mice increased at the same rate for 48 h post exposure to doxycycline, however between 48 h and 72 h in the induced group the parasitaemia decreased to below a detectable level (1×10^4 per ml blood) (Figure 4-5b). This demonstrates that depletion of NEK 12 kinases is fatal to BSF parasites in culture and in mammalian infections.

In order to demonstrate that the NEK 12.1 mRNA was being downregulated by the NEK 12 RNAi construct a qPCR was performed to assess this. Samples of RNAi induced and uninduced cultures were taken at 24 h and used in total RNA extractions, which were then used to make cDNA. When these cDNAs were used as templates in a qPCR a downregulation of the NEK 12.1 transcript by over 50% was detected (Figure 3-7). In order to determine if the ablation of the mRNA was causing depletion of the NEK 12.1 protein an allele of NEK 12.1 was endogenously tagged with a GFP fusion.

To achieve this the NEK 12.1 CDS was amplified and cloned into pENT-6-BG (pGL 1796) to generate an N-terminally tagged GFP-TY-NEK12.1 fusion (Kelly et al.,

2007). This CDS was linearised with a unique AgeI site and transfected into cell line 15356 (containing an RNAi cassette to target the NEK 12 genes) (Kelly et al., 2007). Two clones were identified by anti-GFP western blot as expressing a protein of 79 KDa, the predicted size of the fusion protein. The clone possessing the strongest expression (15356B2) was grown for 24 h in the presence or absence of NEK 12 RNAi induction and then the cells were harvested for western blot analysis. After probing with an anti-GFP antibody and appropriate secondary antibody the GFP signal was undetectable in the induced culture demonstrating the RNAi was effective at depleting NEK 12.1 protein from the cells (Figure 4-5c).

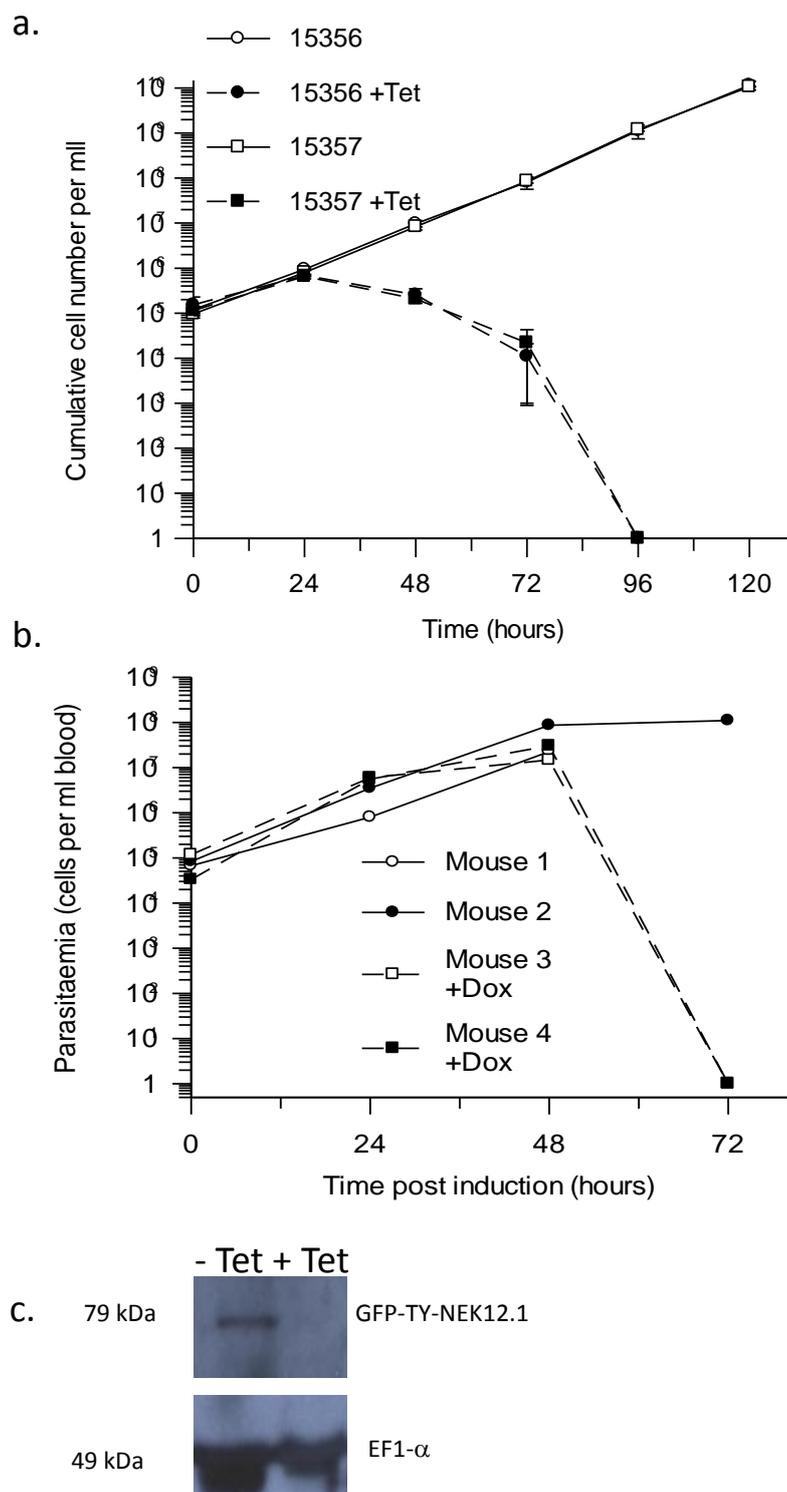


Figure 4-5: Growth curves for *T. brucei* 2T1 BSF RNAi lines (15356/15357) undergoing NEK12 double knockdown, panel a. shows the cumulative cell counts (from haemocytometry) when RNAi is induced in culture. Error bars indicate the standard deviation around the mean of three technical replicates. Cell cultures were maintained between 10⁵ and 10⁶ cells per ml. Panel b. shows the levels of parasitaemia in mice infected with the RNAi line 15356 under induced (+ Dox, doxycycline) and non-induced conditions. Panel c. Western blot against GFP-TY-NEK 12.1 from induced and uninduced cultures of cell line 15356B2, the primary antibody was mouse anti-GFP and the secondary HRP-conjugate anti-mouse. Mouse anti-EF1 α was used to demonstrate equal sample loading.

4.3.3 NEK 12 Phenotype – cell cycle analysis

In order to determine if the depletion of the NEK 12 proteins was causing BSF cell death through a cell cycle defect the DNA content of the cells was analysed by two different methods. Microscopic characterisation of cells following DAPI staining was performed alongside fluorescence activated cell sorting (FACS). In DAPI counts the cells are classified into 1N1K, 1N2K and 2N2K categories - any other configurations can also be scored. If a cell cycle defect is the direct cause of the cell death phenotype this can be expected to appear before the growth defect is apparent. The second method, FACS, uses flow cytometry to determine the DNA content of cells stained with propidium iodide- this method uses far greater numbers of cells than manual counting and removes observer bias. It can detect multinucleated cells but cannot differentiate whether a 2N2K cell is undergoing cytokinesis or not.

To perform manual counts slides were prepared from induced and uninduced cultures and stained with DAPI. The nucleus:kinetoplast configuration of several hundred cells per time point was determined by fluorescence microscopy. At 24 h, before appearance of a growth defect, RNAi of the NEK 12 genes led to an increase in 1N1K cells from 75% to 83% and a reduction in 1N2K cells from 18% to 9% suggesting cells are delaying entry into S-Phase. The number of 2N2K cells is slightly decreased and a small rise in other cell types is noted. Between 24 h and 48 h the number of 1N1K cells drops to 68%, 1N2K to 5% and the 2N2K category increases to 11%. Other cell types rise dramatically to 16%. Other cell types did not consist of a characteristic phenotype. Many contained multiple nuclei with or without kinetoplasts or one nucleus with multiple kinetoplasts. Others contained kinetoplasts or nuclei only. In some cells the nucleolus appeared enlarged. This may suggest aberrant positioning cytokinesis mechanisms in cells depleted of NEK 12 proteins. A clear cell cycle defect is not apparent from this analysis.

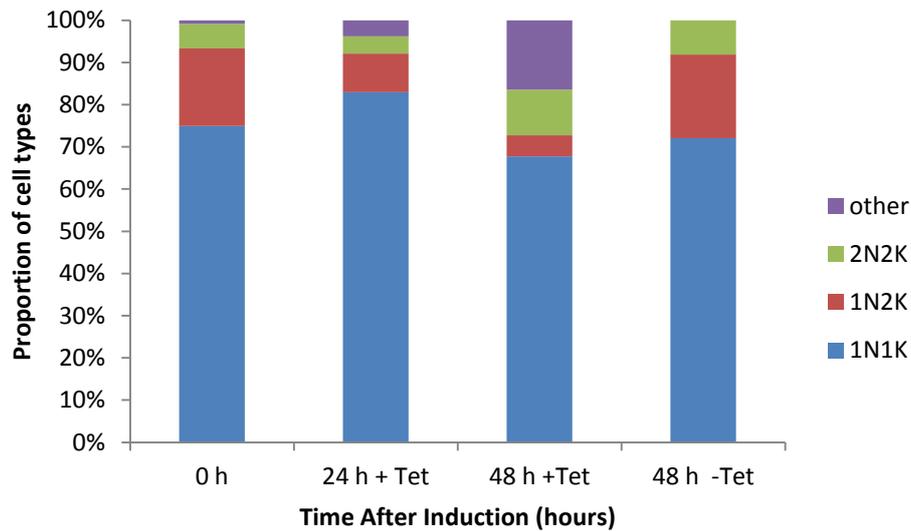


Figure 4-6: Results of DAPI counts showing proportion of cell types in NEK 12 depleted cells (line 15356). N=>200 cells per time point.

When DNA content of the cells was examined using flow cytometry at 24 h a slight increase in 1N1K cells and a slight drop in G2/M cells was observed. At 48 h a larger increase in G1 cells was observed, together with an almost total block of S phase and a reduction in G2/M cells. A very slight increase in sub-G1 cells was observed, possibly indicating the formation of zoids. The increase in “G1” cells may be due to “other” cell types that possess a similar DNA content to 1N1K cells. Again, a clear cell cycle defect is difficult to interpret from these data.

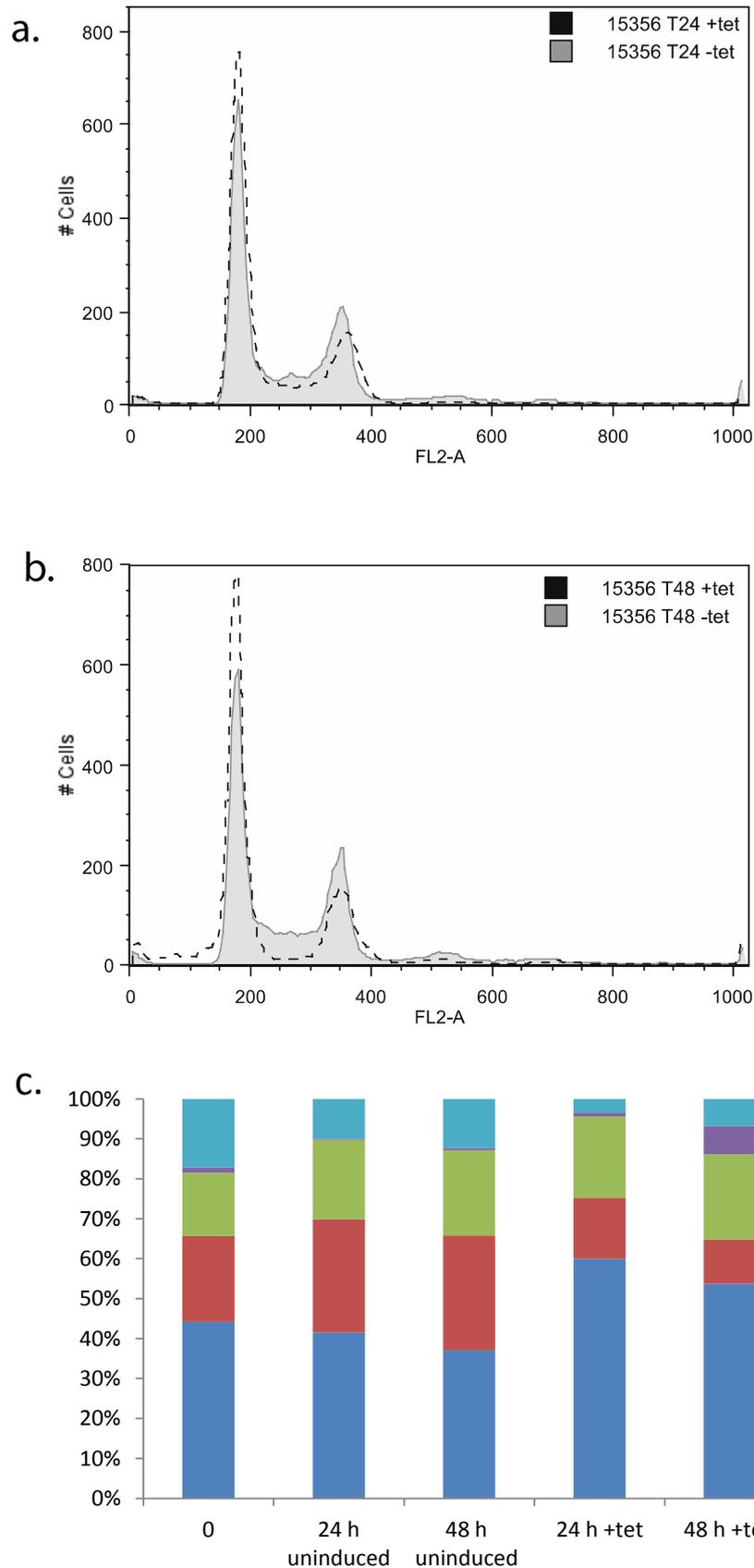


Figure 4-7: Histograms showing the DNA content of NEK 12 RNAi induced (black, dashed line) and uninduced (grey, filled) cells at 24 h (a) and 48 h (b). Cell line used for analysis was 15356. FL2-A peak at 200 corresponds to 1N1K (G1) parasites, and at 400 2N2K (G2/M), intermediate values correspond to S phase parasites. Panel c. Represents the proportion of cells in each phase of the cell cycle as determined by the Dean/Jett/Fox Model using FlowJo software.

4.3.4 Generation of Recombinant NEK 12 Proteins

4.3.4.1 Cloning and mutagenesis

For a cohort of 17 of the smaller NEK (including NEK 12.1) and Orphan kinases the entire CDS was cloned into either the pET-32Xa/LIC or the pET-30Xa/LIC system from Novagen for protein expression experiments. Another cohort of 21 plasmids were generated to express truncations of 13 proteins producing only the protein kinase domain and regions either side intended to allow correct folding and solubility (designed by J. Cowman, University of Edinburgh) (Table 2-6). The pET-30 Xa/LIC vector encodes a C-terminal 6xHis tag (and S-Tag) to allow IMAC purification of the protein; this adds 4.9 kDa to the protein. pET-32 Xa/LIC adds a thioredoxin (Trx) upstream of this with the aim of increasing the solubility of exogenous proteins in the *E. coli* expression system and adds 17.2 kDa to the protein of interest (LaVallie et al., 1993; Hammarström et al., 2006) (Figure 4-8).

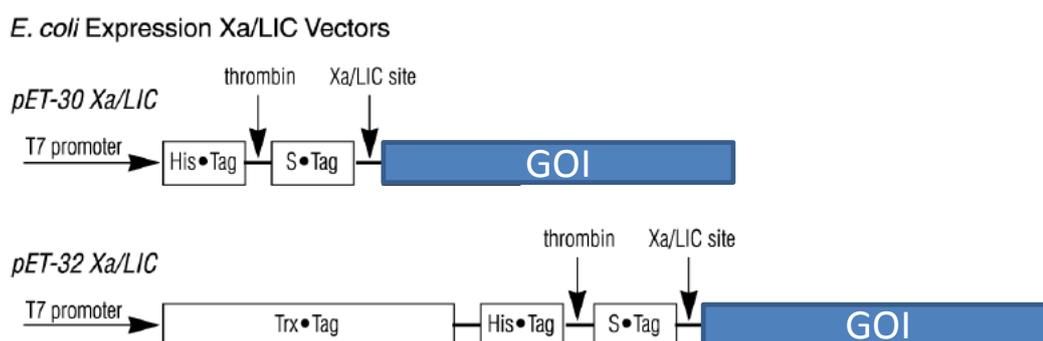


Figure 4-8: Schematic depicting the construct encoded tags fused to the C-terminus of proteins expressed using *E. coli* in this study.

After generating pET-32 Xa/LIC expression constructs for full-length active NEK12.1 and NEK12.2 these were modified by site directed mutagenesis to generate kinase dead mutations of the catalytic lysine to methionine (K70M). These were generated to provide negative controls for kinase activity assays to ensure protein purifications did not contain co-purifying contaminant protein kinases.

4.3.4.2 Cell lines, media, temperature optimisation

Various expression conditions were tested in combination for each construct, varying the cell line, medium and temperature of expression (Table 4-1), at the Edinburgh Protein Production Facility (EPPF). It was apparent that for all the expression constructs tested the optimum cell line to use was Rosetta pLysS in standard LB broth, with expression being induced with 1 mM IPTG at 15°C. Rosetta pLysS contains the pLysS-RARE plasmid to encode for extra copies of rare tRNAs of *E. coli* to increase yields of eukaryotic proteins and to reduce leaky expression of exogenous proteins. An example of expression testing is shown in Figure 4-9 for the proteins in Table 4-2. Other conditions were tested but data are not shown as the conditions were inferior in terms of generating soluble protein. NEK 12.1 was expressed at high levels and was detected in coomassie stained gels as well as the anti-his western blot Figure 4-9. The bands from these gels were excised and analysed by MALDI-MS to ensure the overexpressed protein was the intended target protein; this was the case for all 12 tested at EPPF.

Table 4-1: Variables tested to optimise protein expression,

Cell Type	Medium	Temp. °C
BL21 (DE3)	LB+ 1mM IPTG	15
C41 (DE3)	Autoinduction TB	20
C43 (DE3)	LB+ 1% Glucose, 2mM MgCl, 1mM IPTG	30
Rosetta pLysS		37

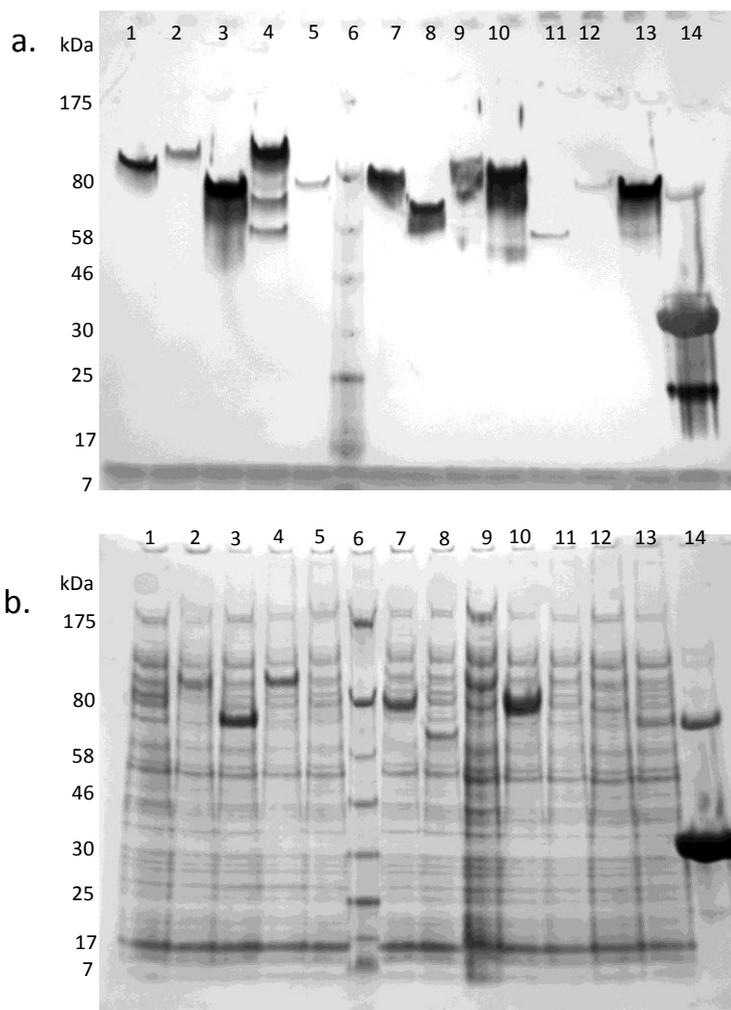


Figure 4-9: Anti-His Western blot (panel a.) showing the expression levels of various NEK and Orphan kinases in the soluble fraction of Rosetta pLysS cells using an overnight induction with 1mM IPTG at 15°C in LB Broth. Panel b. shows a duplicate SDS-PAGE gel stained with Coomassie blue. Descriptions for each lane are presented in Table 4-2.

Table 4-2: Details of lanes in Figure 4-9.

Number	Name	Gene ID	Size+Trx Tag	Plasmid
1	NEK 6	Tb927.3.3080	73.1kDa	pGL 1990
2	NEK 7	Tb927.3.3190	73.3 kDa	pGL 1991
3	NEK 12.1	Tb927.8.7110	66.8 kDa	pGL 1993
4	NEK 14	Tb927.8.1670	82.6 kDa	pGL 1994
5	NEK 15	Tb09.160.0450	68.6 kDa	pGL 1995
6	Ladder	N/A	N/A	N/A
7	NEK 17	Tb927.10.5950	72.1 kDa	pGL 1996
8	NEK 20	Tb11.01.2900	58.8 kDa	pGL 1997
9	NEK 21	Tb11.01.6650	81.6 kDa	pGL 1998
10	NEK 22	Tb927.2.2120	67.8 kDa	pGL 1999
11	Orphan	Tb10.61.0100	55.5 kDa	pGL 2000
12	Orphan	Tb927.3.3290	69.6 kDa	pGL 2001
13	Orphan	Tb11.02.4530	62.3 kDa	pGL 2002
14	His tagged control	N/A	N/A	N/A

4.3.4.3 Purification of NEK12

Various purification strategies were tested at EPPF, University of Edinburgh, to purify NEK 12.1. Initial purification using Immobilised Metal Affinity Chromatography (IMAC) was then followed by ion exchange chromatography or size exclusion chromatography. Various buffers were tested for maintaining NEK 12.1 stability in solution. A simple strategy was found to be suitable for producing NEK 12.1 for activity assays. Pellets of *E. coli* Rosetta pLysS expressing NEK 12.1 were lysed using BPER (Pearce) in the presence of protease inhibitors, the lysates were clarified and TRX-His-NEK 12 was purified using IMAC. Fractions containing NEK 12.1 were buffer exchanged to remove imidazole prior to activity assays. NEK 12.1 was found to be stable in a buffer containing 300 mM NaCl and 50 mM sodium phosphate, pH 8. Protein concentrations of around 0.5 mg/ml could be obtained from this purification strategy allowing the protein to be used in kinase activity assays. Quantities of approximately 10 mg could be extracted from 4 litres of bacterial culture, this protocol was further optimised by workers at EPPF to allow recombinant NEK 12.1 to be further processed for use in crystallization trials. This technique was used with NEK 12.2 but this protein expresses at lower levels than NEK 12.1 and is much more unstable in solution than NEK 12.1, making it much more difficult to work with.

4.3.5 Activity assays/inhibitors of NEKs

Recombinant Trx-His-NEK12.1 and Trx-His-NEK12.2 (purified as described previously) were tested in kinase assays using ^{32}P γ -phosphate radiolabelled ATP to label substrates. Preliminary tests suggested that the NEK12.1 exhibited kinase activity against α -casein but not MBP or histone H1, and displayed an ability to autophosphorylate (Figure 4-10a). Following this discovery, kinase dead mutants (K70M) of NEK 12.1 and NEK 12.2 were generated and purified alongside the wild type protein (using the same conditions) to control for co-purifying, contaminating protein kinases in further activity assays. When the four purifications were used in kinase assays against α -casein (Figure 4-10b) the kinase dead mutant purifications did not exhibit any kinase activity. NEK12.2 exhibited very weak kinase activity against α -casein but the protein appears to be unstable and this reduction in activity may be due to sub-optimal expression, purification or assay conditions. Weak autophosphorylation was detected, though

this was not of the same magnitude as the substrate phosphorylation and was less intense than in preliminary experiments. In order to test if NEK 12.1 could be inhibited by bulky kinase inhibitors, activity assays were conducted in the presence of the bumped PP1 analogue 3MB-PP1 (Figure 4-10c). Due to the instability of NEK 12.2 this was only performed for NEK 12.1. Activity of NEK 12.1 was calculated based on band intensity (measured using ImageJ) of phosphorylated α -casein, and calculated as a percentage of a reaction containing only DMSO. Inhibition of NEK 12.1 kinase activity was observed with an IC₅₀ calculated to be 0.83 μ M with a concentration of radiolabelled ATP at 100 μ M in the assay (Figure 4-10d).

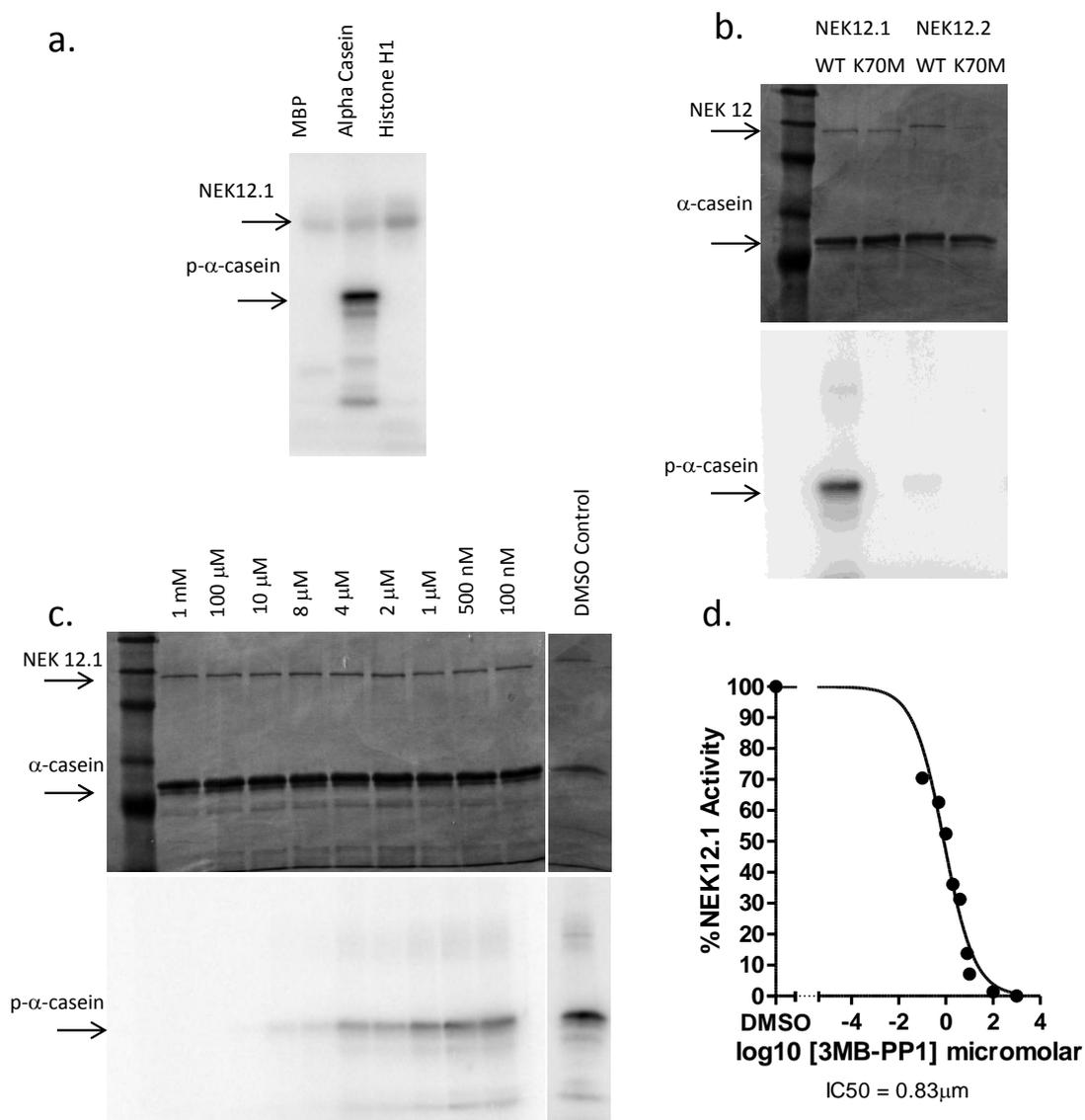


Figure 4-10: Inhibition of NEK12.1 kinase activity with 3-MB-PP1. a. Trx-His NEK 12.1 purifications were tested for kinase activity against three generic substrates. b. These purifications were controlled by the generation of kinase dead mutants (K70M) and Trx-his-NEK 12.2 was also tested for kinase activity. c. Kinase reactions of NEK 12.1 were incubated with increasing concentration of 3-MB-PP1 to determine if NEK 12.1 can be inhibited by a bumped kinase inhibitor. d. The IC₅₀ of 3-MB-PP1 was determined as 0.83 μ M using band densitometry.

4.3.6 Homology Modelling of NEKs and Virtual Screening

Garnering structural information on NEK 12 proteins was deemed important as it would facilitate attempts at rational inhibitor design and chemoinformatic studies to identify potential inhibitors. Despite multiple attempts by collaborators at EPPF, no reliable protocol could be developed that yielded NEK 12.1 crystals that were suitable for X-ray diffraction studies. In lieu of this a homology modelling approach was used to generate limited structural information on the two NEK 12s. NEK 12 sequences were used to search the Protein Data Bank and aligned against several human NEK templates (NEK2: 2JAV, 2W5H and NEK7: 2WQM) using Molecular Operating Environment software. This software was then used to homology model the NEK 12 structure against the crystal structure of the templates. Energy minimisation was used to refine the models, these were then used to observe differences between the active sites of the two proteins. Homology modelling demonstrated a difference in the size of the active site in the NEK 12.1, indicating a larger hydrophobic pocket at the back of the active site due to the presence of a small alanine gatekeeper residue (Figure 4-11).

The models were then used to try and identify potential inhibitors of the two NEK 12s. The pyrrole indolinone ligand in the HsNEK2 structure (PDB ID: 2JAV) was superposed onto the models and used to define the binding site of the NEK 12 models (to define the larger active site of NEK12.1 the ligand was modified by the addition of an ethyl group in PyMol to extend into and define the hydrophobic pocket for data mining). These ligands were used to define sitepoints in the model which allowed the LIDAEUS CAMELSICK 2 library of 4.2 million compounds to be searched for potential ligands. The top 100,000 compounds were then further screened using rigid and flexible docking with the Autodock and Vina software. These were then ranked in order of predicted binding affinity and the top 10 posematched (where rigid and flexible docking produced the same binding pose) were selected for further analysis (shown in Figure 4-12 and Figure 4-13). Of the top 10 hits for NEK 12.1, 6 were predicted to bind in the deep hydrophobic pocket (Figure 4-12). Compounds available from Enamine were purchased for use in further studies (as we had previously purchased compound libraries from this company). These were tested against *T. b. brucei* 2T1 BSF cells in an alamar blue cell viability assay to determine the

IC50 values (Figure 4-12 and Figure 4-13). The .PDB files of the molecular models, the .SDF files of the data-mined ligands and the .PSE files of the docking output are contained in supplemental data file S2 on the appended CD-ROM.

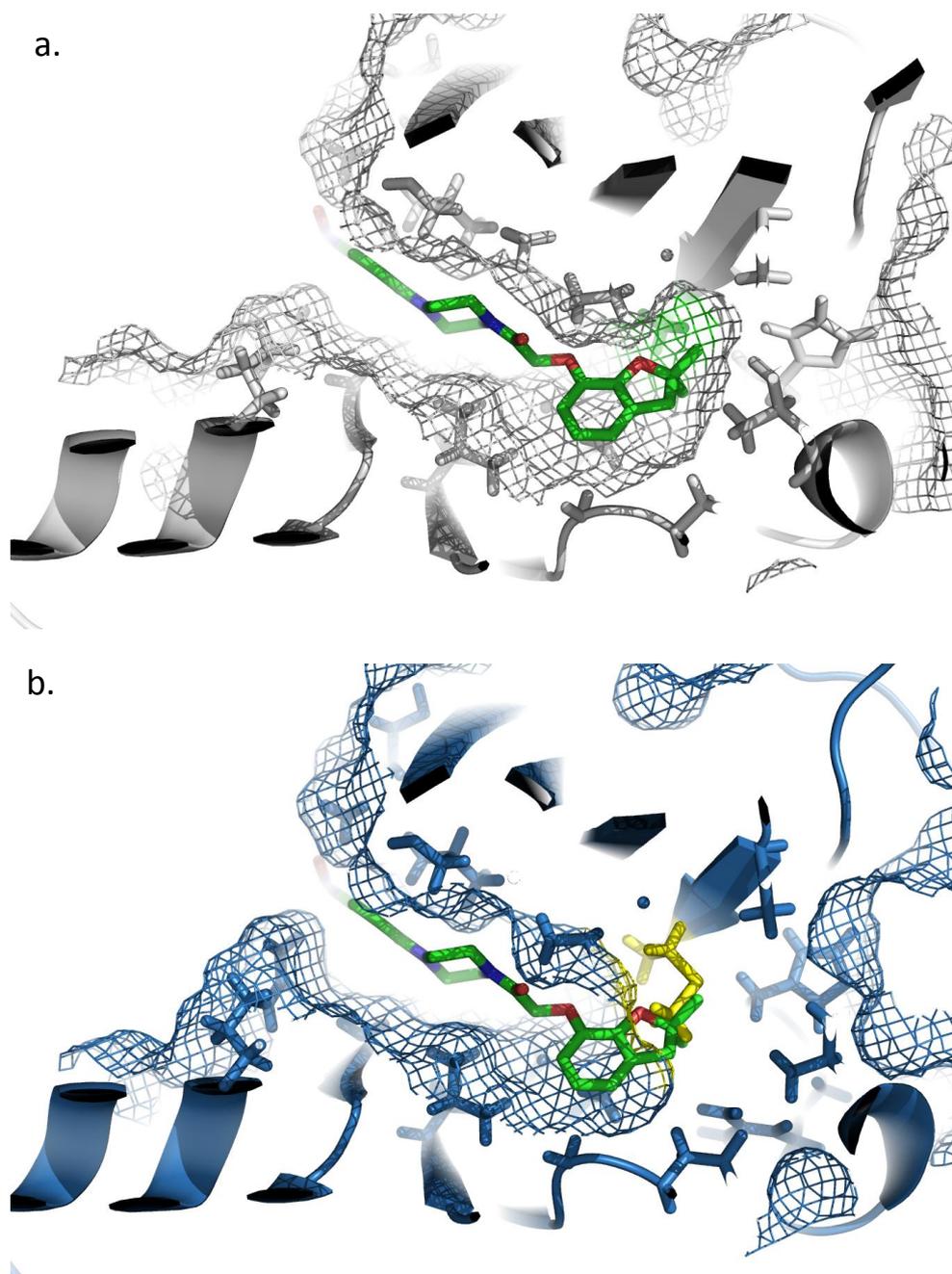


Figure 4-11: Homology models of NEK 12.1 (a) and NEK 12.2 (b) cross sectioned through the active site to show the predicted size of the rear hydrophobic pocket. NEK 12.1 Ala 117 is highlighted in green, NEK 12.2 Met 117 is highlighted in yellow. The ligand (Number 6, Figure 4-12) in the active site was predicted to dock into the NEK 12.1 active site but not NEK 12.2, steric clashes are observed with the Met117 if the ligand is superposed into the NEK 12.2 active site.

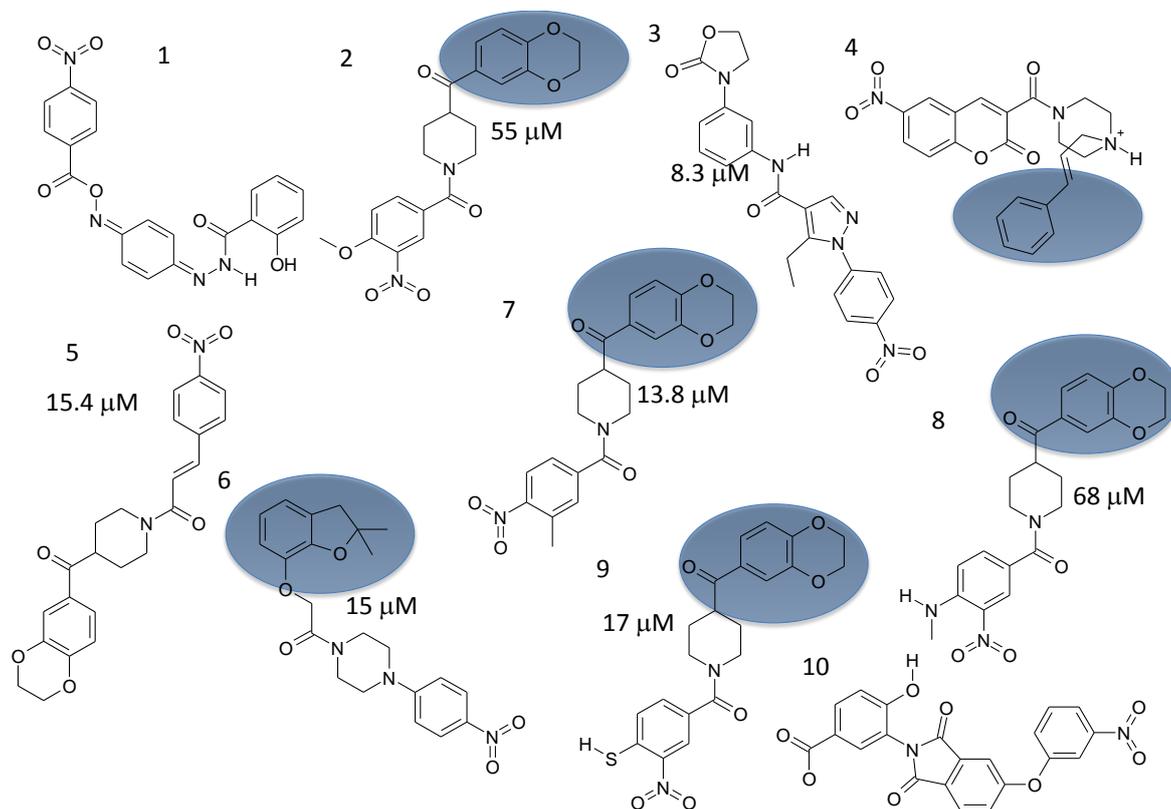


Figure 4-12: Structures of compounds predicted to dock into the NEK 12.1 active site identified by virtual screening. Region of the compound predicted to reside in the deep hydrophobic pocket are highlighted using blue ovals. IC50s are assigned for those compounds purchased from Enamine and tested against BSF parasites.

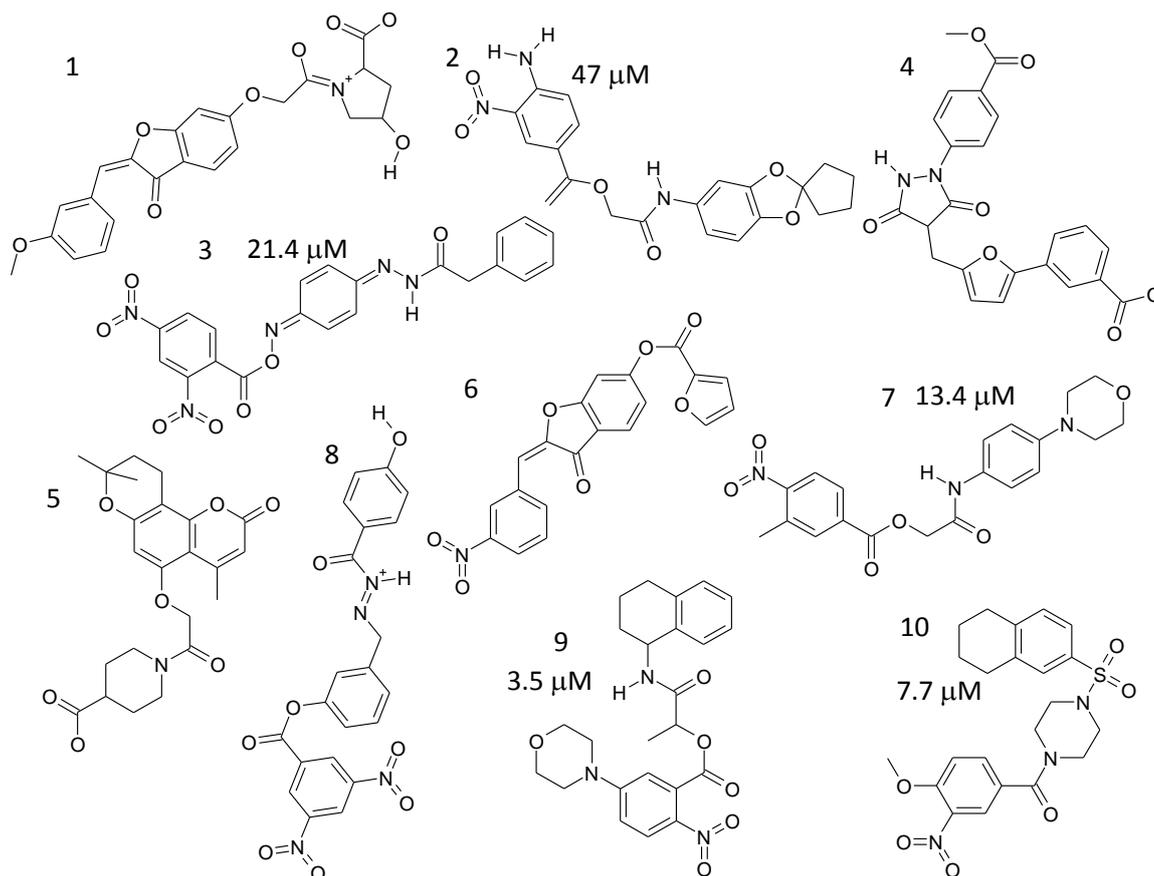


Figure 4-13: Structures of compounds predicted to dock into the NEK 12.2 active site identified by virtual screening. IC50s are assigned for compounds tested against BSF parasites.

4.4 Zinc-Finger Containing Orphan Kinase

4.4.1 Features of ZFO

Only one Orphan kinase was identified as having a detectable growth defect in the kinome-wide RNAi screen. This was the gene Tb09.211.2260, which contains a CDS of 3177 bp and encodes a protein with a predicted size of 118.2 kDa (1058 aa in length). The protein is predicted to contain an N-terminal kinase domain with a long C-terminal extension. This extension is predicted to contain motifs consistent with a C2H2 zinc-finger binding domain. These are classical zinc finger domains where the motif is defined by #-X-C-X(1-5)-C-X3-#-X5-#-X2-H-X(3-6)-[H/C], X can be any amino acid, and numbers in brackets indicate the number of residues. The positions marked # are those that are important for the stable fold of the zinc finger. The final position can be either His or Cys (PFAM, accessed 2012). Zinc finger domains were first identified as DNA binding motifs but have now been shown to bind DNA, RNA, protein or lipid substrates (Klug, 1999; Gamsjaeger et al., 2007; Kröncke and Klotz, 2009). The amino acid

sequence of the putative C2H2 region from ZFK is YTCDA^CGREFASSGALRRH^CSCSC with the key residues highlighted in yellow. Subsequently we have named it Zinc Finger Orphan protein kinase (ZFO).

A number of post-translation modifications have been detected on ZFO (Figure 4-14). Phosphoproteomic data show multiple phosphorylation sites, nine in total (serines 372, 376, 466, 488, and 492 and threonines 374, 468, 470, and 486), in the region between the kinase domain and the zinc finger motif (Nett et al., 2009b). A peptide from ZFO was detected as being palmitoylated in PCF form parasites but the palmitoylation site in the peptide was not defined as it contains several serine and one threonine residue (Emmer et al., 2011).

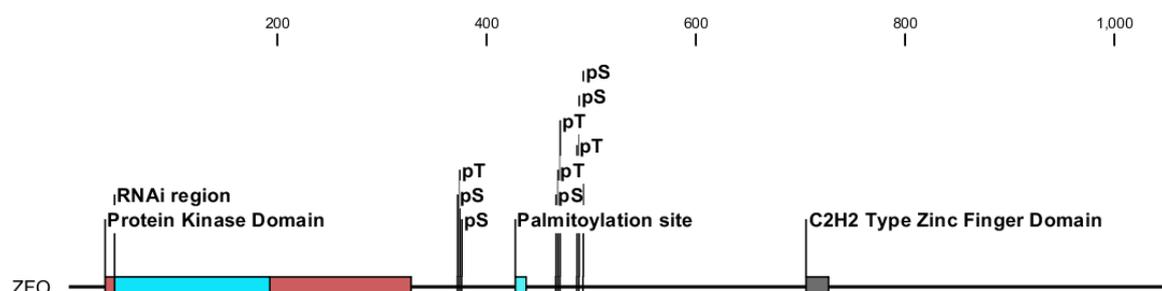


Figure 4-14: Schematic representation of ZFO protein domain architecture with key features annotated. The scale above the diagram shows the amino acid number.

4.4.2 RNAi Analysis

Two cell lines (STL52/STL55) containing an RNAi plasmid (pTL37) targeting ZFO were generated and detected as giving a severe growth defect by alamar blue assay. RNAi ablation of ZFO led to a growth defect becoming apparent after 24 h induction in culture. Between 24 h and 96 h the phenotype became a pronounced cell death defect (Figure 4-15a). The phenotype appeared to be variable between technical replicates in both biological replicates leading to large error margins. This is despite exhibiting consistent alamar blue values, suggesting a variable, pleiotropic phenotype occurs when cells are diluted. When ZFO was ablated from trypanosomes infecting mice the growth defect led to clearance of parasitaemia to below a detectable level by 72 h (Figure 4-15b). Knockdown of mRNA was confirmed by qPCR (Figure 3-7) and further confirmed by western blotting against the product of an epitope-tagged allele of ZFO (Figure 4-15c).

In order to further validate the RNAi knockdown of ZFO the gene was epitope tagged at the endogenous locus to allow western blotting to examine protein levels. This was performed using pHG80, a derivative of pENT-6B-G (Kelly et al., 2007). This allows the endogenous tagging of a gene to create an N-terminal fusion of 3xHA and eGFP. Five hundred nucleotides of the ZFO 5' UTR and a 500 nt fragment of the 5' of the ZFO CDS were cloned into pHG80 between the XbaI and BamHI sites, and were fused together using a unique Ascl site in a three-way ligation. Following confirmatory restriction enzyme digest the plasmid was linearised using Ascl and transfected into STL55 (containing the ZFO RNAi construct). The linearised plasmid inserts into the ZFO locus by homologous recombination, placing a ZFO allele under the control of an aldolase 5' UTR but preserving the endogenous 3' UTR. Five clones were generated from this transfection. Only one clone (STL2266) was observed to express a protein detectable by western blot using anti-HA antibody. A band was observed at 149 kDa, the expected sized of the fusion protein. Following RNAi of the ZFO mRNA for 24 h the fusion protein was ablated to undetectable levels.

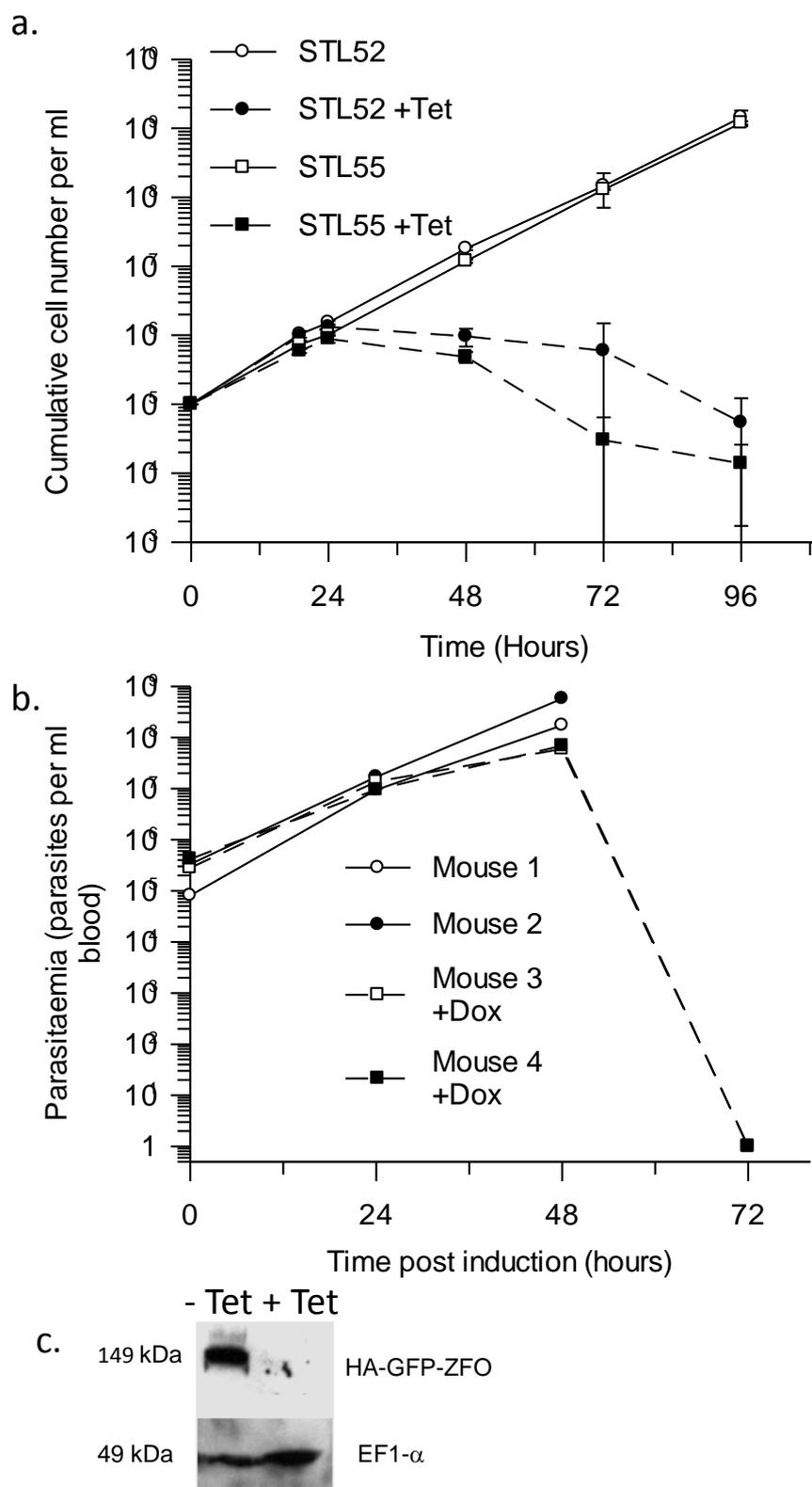


Figure 4-15: Growth curves for *T. brucei* 2T1 BSF RNAi lines (STL52/55) undergoing ZFO knockdown, panel a. shows the cumulative cell counts (from haemocytometry) when RNAi was induced in culture. Error bars indicate the standard deviation around the mean of three technical replicates. Cell cultures were maintained between 10⁵ and 10⁶ cells per ml. Panel b. shows the levels of parasitaemia in mice infected with the RNAi line STL55 under induced (+ Dox) and controlled conditions Panel c. Western blot against GFP-HA-ZFO from induced and uninduced cultures of cell line STL2266, the primary antibody was mouse anti-HA and the secondary HRP-conjugate anti-mouse. Mouse anti-EF1 α was used to demonstrate equal sample loading.

4.5 PDK1

4.5.1 Features of TbPDK1

A predicted AGC family protein kinase, Tb09.160.3480, was identified by BlastP search as having an active site most similar to HsPDK1 (36% identity and 51% similarity). The kinome-wide RNAi screen undertaken previously detected a mild growth defect for this gene. The gene encoding this protein has a CDS of 1155 nt and encodes a 384 aa long protein predicted to have a mass of 43.17 kDa. It is predicted to contain a region C-terminal to the protein kinase domain that is annotated as a FYVE-Zinc Finger domain. This is similar in architecture to HsPDK1 which contains a PH domain C-terminal to the protein kinase domain.

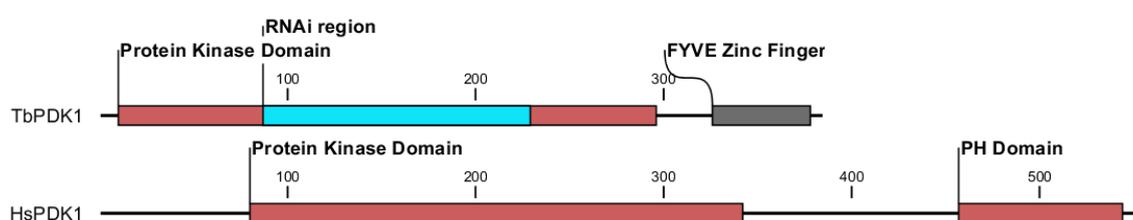


Figure 4-16: Schematic representation of TbPDK1 and HsPDK1 protein domain architecture with key features annotated. The scale above the diagram shows the amino acid number.

4.5.2 RNAi of TbPDK1

Two RNAi lines were generated to investigate the function of the PDK1 homologue (15059/15060). RNAi ablation of TbPDK1 in culture led to a slow onset phenotype with a slow growth phenotype becoming apparent in the cells after 48 h of RNAi induction (Figure 4-17a). Following this, growth became static by 120 h. Cells looked abnormal, becoming enlarged and possessing multiple flagella. Cells exhibited abnormal nucleus and kinetoplast numbers, though this was not formally quantified. When RNAi of TbPDK1 was induced in trypanosomes infecting mice the phenotype became more pronounced with parasitaemia being undetectable after 72 h of RNAi induction (Figure 4-17b). RNAi knockdown of the target mRNA was confirmed by qPCR (Figure 3-7).

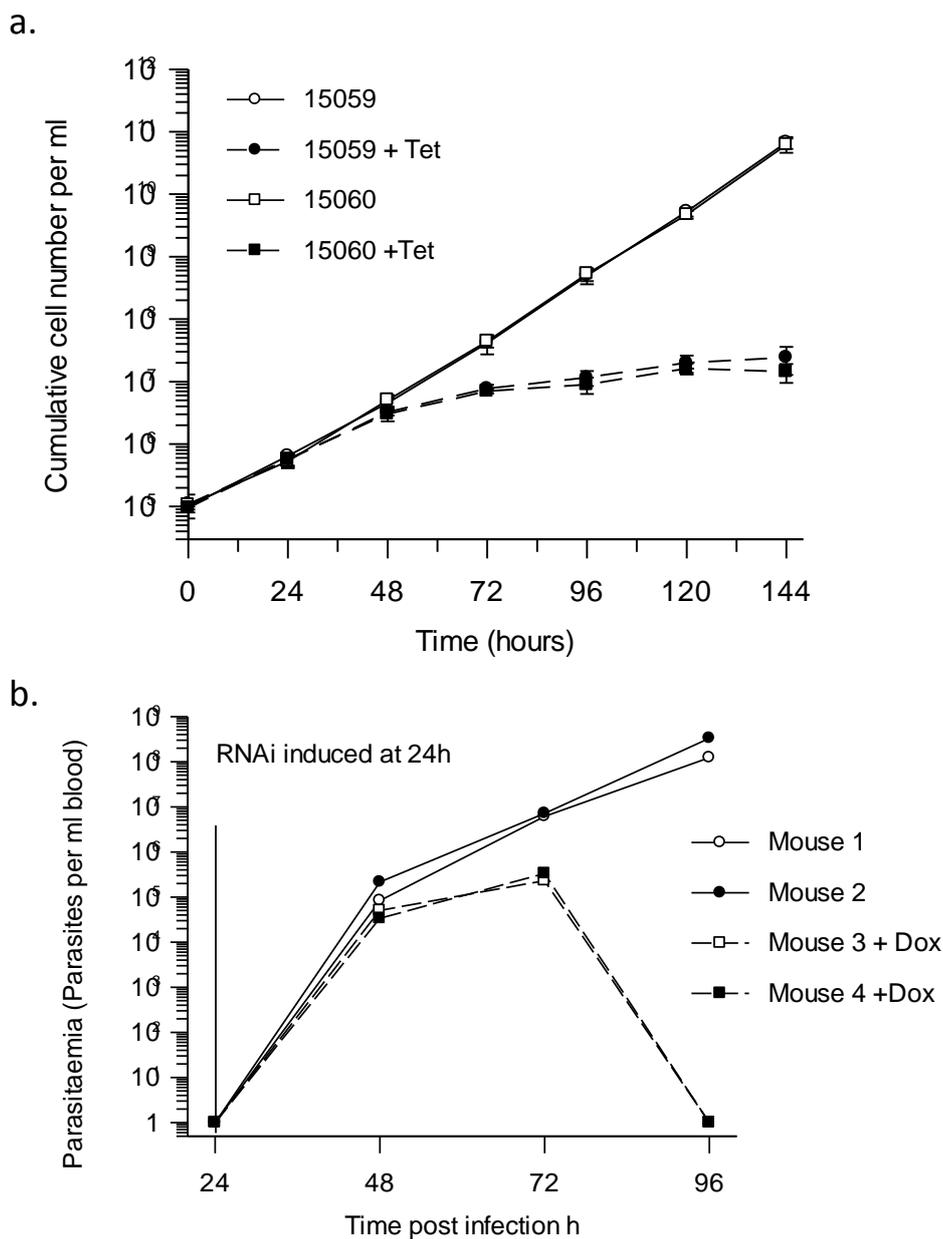


Figure 4-17: Growth curves for *T. brucei* 2T1 BSF RNAi lines (15059/15060) undergoing TbPDK1 knockdown, panel a. shows the cumulative cell counts (from haemocytometry) when RNAi is induced in culture. Error bars indicate the standard deviation around the mean of three technical replicates. Cell cultures were maintained between 10⁵ and 10⁶ cells per ml. Panel b. shows the levels of parasitaemia in mice infected with the RNAi line 15059 under induced (+ Dox) and uninduced conditions.

4.6 Discussion

4.6.1 NEK kinases as drug targets

4.6.1.1 Overview

In the study of the NEK 12 protein kinases their (combined) essentiality to the parasite has been demonstrated by RNA interference studies. Rapid cell death phenotypes were observed in culture and replicated in a more biologically relevant mouse infection. The knockdown of the target mRNA and protein was confirmed by two methods, qPCR and western blotting (of an epitope tagged NEK12.1). Abolition of the two NEK 12 protein kinases did not lead to an characteristic morphological or cell cycle defect. Cytokinesis appeared to continue, though correct segregation of nuclear DNA or kDNA appeared severely impaired leading to aberrant NK configurations. The NEK 12 proteins were homology modelled and this suggested that the small alanine gatekeeper residue of NEK 12.1 allowed the active site to expand, creating a deep hydrophobic pocket - a feature considered amenable for engineering specificity in potential inhibitory compounds. NEK 12.1 recombinant protein was amenable to expression in a bacterial system and could be purified to homogeneity yielding stable protein suitable for activity assays. Protein kinase activity assays, controlled by appropriate kinase dead mutants, demonstrated the recombinant NEK 12.1 was active and had a substrate preference for α -casein. These assays could be inhibited by the addition of 3-MB-PP1, a protein kinase inhibitor with a bulky side group that should only inhibit protein kinases with a large hydrophobic back pocket (Bishop et al., 2000a, 2000b). This demonstrates NEK 12.1 is a suitable candidate for inhibition with more selective bulky inhibitors, though screening to identify the optimum inhibitor is probably required. These findings indicate that NEK 12.1 is potentially a very attractive molecule from the viewpoint of rational drug design. Some hurdles must be overcome to consider it a validated drug target and must be investigated further.

4.6.1.2 Genetic Validation

The first hurdle is to determine if NEK 12.1 is an essential kinase. The current knockdown construct targets two closely related genes leaving the question; is either NEK 12 essential to BSF trypanosomes in an individual knockdown? In order

to answer this it was attempted to design RNAi sequence to target the 3' UTRs of each gene individually, as they are more diverged from each other than the relative CDS. However, when used in an RNAi query the UTRs were too similar to other regions of the trypanosome genome, and would have led to downregulation of non-target mRNAs making this strategy impossible. Other strategies to validate these genes included gene knockout, which could have been used to show if either gene was not essential i.e. it can be knocked out. However, if a gene cannot be knocked out it only suggests essentiality, as the failure to stably transfect knockout constructs may be due to transfection failure. An inducible knockout system based on Cre-recombinase has been developed for use in trypanosomes that could allow inducible knockout of essential genes but is not in widespread use due to concerns over Cre toxicity (Scahill et al., 2008). Other techniques exist in other parasitic organisms such as plasmid shuttling or fusing the gene of interest to an inducible destabilisation domain, but these have not been established in *T. brucei* (Herm-Götz et al., 2007; Morales et al., 2010). In order to attempt to validate the genes individually the technique we have chosen is a re-expression of RNAi refractory mutants to complement either wild-type gene that is ablated by double RNAi of both NEK 12s.

4.6.1.3 Re-Expression of RNAi Refractory NEK 12 Mutants

The data presented only go so far to validate the NEK 12 kinases as drug targets. Due to the nature of the double-knockdown achieved with the current RNAi system the relevance of each individual NEK 12 is unclear. In order to resolve this and define the more clearly the role of each NEK 12 we aimed to perform re-expression experiments where a recoded, RNAi refractory gene is expressed to complement the depletion of a wild type allele by RNAi. By doing this for NEK 12.1 and NEK 12.2 the relative contribution to the double RNAi phenotype could be determined. For example it may be that both NEK 12s are essential individually, one is individually essential, or neither are individually essential. The ability to make functional mutants, e.g K70M kinase dead mutants, would also allow the validation of whether NEK 12 kinase activity is essential for BSF trypanosomes. The precedent for re-expressing RNAi refractory mutants in an RNAi background (in *T. brucei*) was set by Rusconi *et al.* 2005 by using targeting UTRs of genes and reexpressing them under the control of other UTRs (Rusconi

et al., 2005). Szöör *et al.* 2010 used a different approach by recoding the CDS (Szöör et al., 2010).

In order to re-express RNAi refractory mutants of NEK 12.1 and NEK 12.2, and to express kinase dead mutants *in vivo* the expression vector pRM481 was recoded to include more user-friendly restriction sites and AttR Gateway Recombination sites (Proudfoot and McCulloch, 2005). Gateway sites were included in an attempt to make this system a tool suitable for medium-throughput cloning of these vectors to facilitate future functional genetic studies and to remove any problems associated with conflicting restriction sites. It was also engineered to contain a C-terminal 6xHA epitope tag to allow detection of recoded protein expression. To generate RNAi refractory NEK 12 genes their amino acid sequence were reverse translated, using *T. brucei* 927 codon bias, and manually assessed to ensure no regions of identity to wild-type genes over 20 nt in length. The recoded sequence was then used in an RNAi search to ensure no homology to the wild-type gene was detected and thus would not be downregulated by the NEK 12 RNAi target sequence (Redmond et al., 2003). The synthetic constructs were delivered and cloning attempted, though problems with the efficiency and accuracy of Gateway reactions have limited the usefulness of this approach so far.

The advantage of re-expressing RNAi refractory mutants is the ability to mutate key features of the re-expressed protein to assess the effect *in vivo* once the wild type protein has been degraded. This should allow the importance of NEK 12 activity to be tested and determine whether the observed phenotype is due to the lack of NEK 12 activity or lack of NEK 12 protein. Should NEK 12 form part of a multiprotein complex, ablation of NEK 12 may disrupt the complex and lead to a more severe phenotype than, for example, replacing NEK 12 with a kinase dead mutant. The importance of NEK 12 protein kinase activity to the parasite is an important question to answer; all current protein kinase inhibitors target catalytically competent protein kinases. Inactive or scaffold proteins related to protein kinases may present poor drug targets as novel inhibitors would have to be considered, rather than those that target the active site. Some proteins, such as Kinase Suppressor of Raf (KSR), have structural features suggestive of a protein kinase, though are catalytically inactive (Roy et al., 2002; Clapéron and Therrien, 2007). It acts as scaffold for the mammalian MAPK pathway. These may

have evolved from protein kinases to become catalytically inactive MAPK pathway supports. As such, proteins like KSR may present limited potential as drug targets due to their lack of enzymatic activity and the need for potential inhibitors to disrupt protein-protein interactions (Brennan et al., 2011). This is due to the fact that disrupting protein-protein interactions, which take place across a large surface area, may be more difficult than inhibiting a defined active site. However, this is being attempted for some CDK:Cyclin interactions (Kontopidis et al., 2003; Andrews et al., 2006). By re-expressing kinase active and dead mutants of each NEK 12 protein it should be possible to test if kinase activity is required for each protein for its function *in vivo* and if either NEK 12 is individually essential to *T. brucei*. In order to fully genetically validate NEK 12.1 as a drug target then these two requirements must be satisfied.

4.6.1.4 Chemical Validation

Should it be the case that NEK 12.1 is an essential, active protein kinase in the trypanosome (which is suggested by the kinase activity of the recombinant protein) then the next step is to screen candidate protein kinase inhibitors. This can be performed using a multitude of techniques. In this study the results of a virtual screen have been presented. Virtual screening has the advantage of being a “dry” technique; therefore is cheap, fast and relatively simple. The disadvantage is that it may not represent the real biological interaction of a receptor and ligand and since many different techniques exist, and each may produce a different result, standardisation is a problem. However, it is considered by some to be good first-pass screening technique to identify potential inhibitors or compound types/families for further investigation (Schneider, 2010; Scior et al., 2012).

Ideally the virtual screen of the NEK 12s would have been conducted against a structure of NEK 12 determined by X-ray crystallography. In the absence of this, a homology model was used. Although homology models can be viewed as poor substitutes for real structural data, protein kinases contain many conserved amino acids vital for function, allowing good alignments of the query to the template. The template used in this model also had an inhibitor in the active site, directing residues towards it and increasing rigidity in this region. This should increase reliability of the model in this region. An obvious difference was

demonstrated between the back hydrophobic pocket of the two enzymes, with NEK 12.1 being much larger and canted upwards from the rest of the active site. The main deficiency in the models was the modelling of the activation loop. The template used to generate the models did not contain information on the activation loop due to its high flexibility producing poor X-ray diffraction patterns. This led to potentially very poor modelling of this region, however none of the ligands are predicted to interact with this area.

The results of a virtual screen identified a diverse range of compounds predicted to bind into the active sites of both enzymes, including compounds with bulky ring structures predicted to interact with the rear hydrophobic pocket of NEK 12.1. Many of the compounds contain a nitro group that is predicted to occupy a solvent accessible site. A number of compounds identified from this screen were tested for their ability to inhibit the growth of BSF trypanosomes and several of those predicted to bind NEK 12.2 gave IC₅₀s in the low micromolar range. Compounds predicted to bind in the deep hydrophobic of NEK 12.1 actually had higher IC₅₀s. Their ability to inhibit recombinant NEK 12.1 has not yet been tested but doing so might yield information on the structure-activity relationship (SAR). The reasons for the bulky, predicted NEK 12.1 inhibitors not having much effect on the BSF parasites may be due to solubility or cell permeability issues or NEK 12.1 just not being essential. As the inhibition profile of these compounds is unknown they may well be generating off target effects. However, even if this is the case, it may not necessarily rule them out as potential lead compounds - if they inhibit a range of trypanosome NEK kinases (but not human kinase) this could exploit the concept of polypharmacology making them more toxic to the parasite than a very specific inhibitor (Karaman et al., 2008; Ghoreschi et al., 2009). It is clear that *in vitro* bulky inhibitors such as 3-MB-PP1 can inhibit NEK 12.1 and that it has an inhibitory effect on parasite growth with an IC₅₀ of 0.83 μ M. Profiling these inhibitors using chemical proteomics *in vivo* may provide more realistic IC₅₀s for NEK and other kinases - NEK 12.1 and 12.2 were amongst 7 NEK kinase detected by a recent chemical proteomics screen investigating inhibitors of GSK3 and the NDR kinase PK50/PK53 (Urbaniak et al., 2012b).

Other techniques to search for inhibitors of NEK 12.1 could include using the recombinant protein in a surface plasmon resonance screen (SPR) of a fragment

library or to use the recombinant protein to develop a high-throughput screen (HTS) based on measuring enzyme activity. Fragment screening using techniques such as SPR can be used to detect very weak interactions of a ligand with a receptor and also use minute quantities of material. It allows the identification of fragments from drug-like molecules, which can then be combined to design specific and potent inhibitors of the protein of interest (Neumann et al., 2007).

In order to set up a HTS against large compound libraries the activity assay used to measure the protein kinase activity must be re-formatted. It would be unsuitable to use a radioactive assay for both practical and safety concerns, and a non-radioactive assay such as an ADP/Kinase-Glo assay or IMAP fluorescent polarisation assay would be preferable (Tanega et al., 2009; Walker et al., 2011c). This would require identification of a suitable peptide substrate for NEK 12.1 (whose *in vivo* substrate is unknown), which can be carried out by screening a peptide library. Some human NEKs such as NEK 6 (Lizcano et al., 2002) and NEK 2 (Cell Signalling Technologies) have been profiled for substrate preference and this may present a logical place to start looking for a NEK 12 peptide substrate.

4.6.1.5 Biological Role of NEK 12 protein kinases

Though the importance of the NEK 12 protein kinases to BSF *T. brucei* was demonstrated in this study, the biological role was not well investigated and one can only speculate as to their role in the cell. NEKs have been determined to play diverse roles in other organisms, though there are recurring themes. It appears that NEKs regulate cytoskeletal events, ranging from the development of cilia and axons to the control of mitotic spindle segregation. Trypanosomes have highly dynamic microtubule-base structures and organelles and accurate control of these is crucial to ensuring accurate morphology, motility and cell division (Benz et al., 2012; May et al., 2012).

In the kinome-wide RNAi screen only 3 NEKs were detected as being important to BSF parasites after 3 days of RNAi ablation. Data from BSF RIT-Seq analysis confirmed our phenotypes for the NEK 12 kinases and NEK 16 (Tb927.10.14420), though RIT-Seq showed a stronger phenotype for NEK 16. RIT-Seq also detected Tb11.01.2900 (NEK 20) as generating a loss of fitness phenotype after 6 days RNAi, which the pTL screen would not detect (Alsford et al., 2011). The chance

of false negatives existing in these data is low after assessment of the library, therefore it appears that few of the NEKs are individually essential to BSF parasites. Several possibilities could explain this. The first is that the NEKs are so important to the parasite that it has evolved a battery of genes to enable redundancy in the system. If NEKs of *T. brucei* operate in cascades as seen in mammalian NEKS (Belham et al., 2003) then other protein kinases may be able to step in and complement the loss of another. As several NEKs are possessed in an array (Tb927.10.5930, Tb927.10.5940, Tb927.10.5950) this may suggest that the parasite is attempting to increase its gene dose of these protein kinases, further hinting at their importance. The role of redundancy in these systems can be probed by conducting double RNAi experiments or by attempting a chemical genetics approach. This could combine RNAi lines not exhibiting a growth defect combined with application of libraries of potential protein kinase inhibitors that also do not generate a lethal phenotype. Observation of the cells for synergistic effects of RNAi plus inhibitor can be used to identify different protein kinases that are not individually essential but together generate a synthetic lethal. This approach may start to allow the assembly of signalling pathways in *T. brucei*. A library of non-lethal inhibitor-like molecules should be identified from screens of a 30,000 compound library made available by GSK under the Open Lab Foundation (<http://www.openlabfoundation.org/NEU-CSIC.aspx>) and these may provide a useful resource in any future synthetic lethal screen.

An alternative is that the NEKs examined in BSF form really are not essential for normal replicative growth in culture. However, if these proteins are involved in the complex remodelling events that occur during the other life cycle stages then they may show essential functions in these stages. TbMKK1 has been shown to be dispensable in BSF parasites but absolutely essential for progressing the parasite through the tsetse vector, potentially at the proventricular or epimastigote forms (Morand et al., 2012). The NEK kinases of *Plasmodium*, and *Giardia* have been shown to perform function at various different stages of development and it may be that in trypanosomes that NEKs are temporally regulated during different stages of the life cycle and thus an RNAi phenotype would only be observed in a given stage or during stage transition (differentiation). As the 2T1 strain is monomorphic, and any protein kinase important to stumpy BSF parasites may not be detected in this screen, a role in

slender to stumpy differentiation cannot be excluded from the potential functions of NEK kinases. Data from RIT-Seq experiments show that Tb927.2.2120 (NEK 22) and Tb927.4.5390 (NRKA) are essential in differentiating (BSF-PCF) trypanosomes and that Tb11.01.2900 (NEK 20) and Tb927.8.1690 (NEK 14) are essential to established procyclic form trypanosomes (Alsford et al., 2011). These experiments observe only part of the *T. brucei* lifecycle and important roles for NEK kinases are likely to exist in the other developmental stages found in the tsetse fly.

One potential role of NEK kinases that could be investigated is regulating the disassembly of the nuclear pore complex (NPC) since various NEKs are involved in the regulation of this during mitosis in mammalian (HeLa) cells. HsNEK 6 and HsNEK 7 were shown to phosphorylate the nucleoporin Nup98, a protein which is predicted to be conserved in trypanosomes (DeGrasse et al., 2009; Laurell et al., 2011). Performing a blast search of the *T. brucei* proteome with these two NEK sequences returns no clear homologues, with Tb927.3.3190 the top hit in both queries, though an RNAi phenotype was not detected for this NEK kinase during the pTL screen. The role of NEKs in regulating the flagellum has not been investigated in *T. brucei* though it may be an interesting avenue for investigation. The flagellum of kinetoplastids is a complex organelle responsible for multiple functions such as motility, attachment and potentially signalling. Several protein kinases are known to regulate flagellum length in *Leishmania* promastigotes (Bengs et al., 2005; Erdmann et al., 2006), and cilia formation in *Tetrahymena* and higher eukaryotes (Mahjoub et al., 2005; Wloga et al., 2006). More detailed phenotypic analysis of NEK RNAi lines could measure the flagellum for any changes in length or morphology.

An interesting feature of NEK 12.1 and NEK 12.2 is the PH Domain. 7 NEK kinases are predicted to possess this domain (Parsons et al., 2005b), which may regulate the localisation of this protein kinase or direct it to its substrate (Lemmon et al., 2002; Lemmon, 2004). This domain was originally identified as binding to phosphoinositides. However, the majority of PH domains described do not independently localise to phospholipids and binding to small GTPases is another proposed target of the PH domain (Lemmon, 2004). Identifying where these proteins localise could provide information on their function. Also, if the PH domains do interact with other proteins then co-immunoprecipitation

experiments should be able to detect this, allowing better characterisation of their biological role. Localisation experiments to probe the function of NEK 12.1 were considered. The localisation of the protein within the cell is unknown. Defining this could provide some information on its potential function. Recombinant NEK 12.1 (cleaved of its Trx and His tag) was used to immunise chickens; subsequently the IgY from the yolks of these chicken's eggs was purified. Even after affinity purification of this IgY against recombinant NEK 12.1 it has proved impossible to generate a clear, specific signal in western blots against parasite lysates. If this resource had been available it would have allowed a number of localisation experiments to be performed. It may also have allowed immunoprecipitation experiments to attempt to identify potential interacting proteins of NEK 12.1 - especially as the PH domain found in NEK 12.1 may facilitate protein-protein interactions.

Another feature of NEK 12.1, its small gatekeeper residue, may present opportunities to probe the function of protein kinase. If a selective inhibitor can be found for NEK 12.1 it could allow chemical genetic studies to dissect the function of NEK 12.1 over NEK 12.2. This approach has previously been used to investigate the function of TbAUK1 (the homologue of Aurora kinase A) in procyclic form parasites (Li et al., 2009). The authors suspected a dual role for TbAuk1 in mitosis and cytokinesis but could not perform RNAi in synchronous PCF cultures and observe both phenotypes clearly. The highly selective Aurora kinase inhibitor VX-680 was found to inhibit TbAuk1 *in vitro*, phenocopying the RNAi and overexpression of kinase dead mutant phenotypes. Using this inhibitor on synchronised PCF cells allowed the elucidation of TbAuk1s multiple roles in the cell cycle. This demonstrates the power of highly selective chemical inhibitors over genetic techniques.

One way specific inhibitors could be used to investigate NEK 12.1 is in trying to identify its substrate. Previously, protein kinases with large gatekeepers have been engineered to contain smaller gatekeepers - which are compatible with bulky inhibitors. This allows specific inhibition of the so called analogue sensitive kinase allele (ASKA) (Bishop and Shokat, 1999; Bishop et al., 2000b). Libraries of bulky inhibitors have been developed to allow optimum inhibition of the ASKA protein kinase. The mutation often introduced is to change the gatekeeper to a glycine or an alanine. This approach is being used to study TbPLK (de

Graffenried, Unpublished). As NEK 12.1 already contains this arrangement it may be able to take advantage of the existing material (compound series) available for ASKA studies.

The IC₅₀ determination of 3-MB-PP1 against NEK 12.1 demonstrates that the small, alanine gatekeeper residue of NEK12.1 allows the accommodation of a bumped inhibitor. The optimisation of a kinase assay and further characterisation of NEK 12.1 would be crucial prior to making finer assessment of inhibitor potency. NEK12.1 could be characterised further by identifying the k_m and v_{max} , optimum ATP and substrate concentrations (processes that would also need to be conducted before developing an HTS assay). However, the ability for NEK 12.1 to accommodate bulky compounds could allow specific inhibition of NEK 12.1 by compounds that are only suitable as research tools. This could allow more accurate definition of the phenotype of blocking NEK 12.1 activity. If a specific inhibitor existed for research purposes then comparative phosphoproteomics could be used to compare inhibitor treated cells (or cell lystate) to untreated controls and look for changes in the patterns of phosphorylation, therefore identifying NEK 12 substrates or pathways. The large active site also opens the possibility of using bumped ATP analogues that allow radiolabelling or thiophosphorylation of the substrate followed by affinity purification and mass spectrometry to identify the substrate of NEK 12.1 (Johnson and Hunter, 2005; Elphick et al., 2007; Chi and Clurman, 2010).

In terms of *in vivo* analysis with a bulky inhibitor, the other bulky derivatives of PP1 could be screened for effects against the recombinant NEK 12.1 kinase and any potent inhibitors tested against the parasite. This may identify inhibitors that phenocopy the RNAi of NEK12 or generate a different phenotype due to the lack of NEK 12.2 inhibition (assuming that NEK12.1 is the only ePK capable of being inhibited by bumped inhibitors). 3-MB-PP1 was tested against 2T1 BSF cells, generating an IC₅₀ of 13.7 μ M. Though perhaps not ideal, relatively high concentrations (5 μ M) of bumped inhibitors including 3-MB-PP1 have been used to study calcium-dependent protein kinase 1 mediated egress of *T. gondii* from host cells (Lourido et al., 2010). Screening of the other derivatives of PP1 may reveal more potent inhibitors. Also, one other protein kinase (Tb927.10.3340) possesses an alanine gatekeeper. When ablated by RNAi this did not generate a

detectable phenotype so the effect of bulky inhibitors on the parasite may be unclear unless specificity is engineered into other parts of an inhibitor.

This finding has implications for ongoing work to establish ASKA versions of other trypanosome protein kinases such as TbPLK. Mutations made to the gatekeeper of PLK reduced its activity and a compensatory mutation was made in the N-terminal lobe, allowing inhibition of sensitised-TbPLK with 3-MB-PP1 (de Graffenried, Unpublished). However, if 3-MB-PP1 also interferes with NEK 12.1 then it could complicate the findings of any TbPLK ASKA assay. Recently, a new way of targeting sensitised protein kinases with the use of an engineered cysteine gatekeeper residue was demonstrated. Modifying the gatekeeper to cysteine allows it be targeted by electrophilic inhibitors which may present a much more specific way of doing these experiments as cysteine gatekeepers are particularly rare (Garske et al., 2011). In addition, these inhibitors form a covalent bond to the cysteine and therefore irreversible inhibition of the engineered protein kinase.

A word of caution must be said for the gatekeeper residue - point mutation of this single amino acid can lead to drug resistance in a given kinase. The most pertinent example of this is resistance of the Abl kinase to imatinib, dasatinib and nilotinib, caused by mutations of the small threonine gatekeeper (often to an isoleucine) (Barouch-Bentov and Sauer, 2011). In the hypothetical situation where a NEK 12.1 inhibitor, taking advantage of its small gatekeeper residue, entered clinical use then drug resistance would potentially be easily evolved.

4.6.2 ZFO and PDK1 as Drug Targets.

The data presented suggest that both ZFO and PDK1 might be genetically validated drug targets, though requirement for activity *in vivo* must be confirmed. The Orphan kinases do not fall into any classical families of protein kinase, therefore gaining structural data on ZFO would be interesting from a drug development and basic research viewpoint. The characterisation of ZFO could be very interesting due to the zinc-finger binding region. These are stereotypically involved in binding to nucleic acids- trypanosomes regulate gene expression post-transcriptionally and RNA binding proteins are key to this. Zinc-finger containing proteins (including protein kinases) have been previously

identified as having important roles in *T. brucei* (Vassella et al., 2001; Kramer et al., 2010; Benz et al., 2011; Ouna et al., 2012). Analysis of ZFO should include localisation studies, and immunoprecipitation studies of the HA tagged allele in order to attempt to identify interacting proteins or nucleic acids.

PDK1 in mammals is a well-characterised enzyme and is the focus of much research. It is unclear whether the protein kinase investigated here is a functional homologue of HsPDK1. Trypanosomes are predicted to possess numerous conventional and unconventional PI-Kinases and are susceptible to inhibitors of PI3Kinases (Bahia et al., 2009; Diaz-Gonzalez et al., 2011). A blast search of the trypanosome kinome with HsPKB protein sequence identified a protein kinase with similarities to PKB (Tb927.6.2450), including an N-terminal PH domain, suggesting there may be enzymes and pathways analogous to some of those involving PDK1 in mammals. The FYVE type zinc finger motif, predicted in TbPDK1, is known to bind PI3 and localisation of this protein should also be conducted (Kutateladze and Overduin, 2001; Tibbetts et al., 2004).

In terms of chemical validation an attempt to target TbPDK1 (*in vivo*) was undertaken. In collaboration with Dr Rudi Marquez at the University of Glasgow Department of Chemistry, over 30 compounds (based on a “a polysubstituted pyridine framework”) developed as potential inhibitors of human PDK1 were assayed for activity against *T. brucei* 427 wt BSF cells using an alamar blue assay. Disappointingly, all compounds bar one had an IC₅₀ above 75µM; Compound “148” had an IC₅₀ of 19.35µM (data from N. Kasumu, MSc), which could be due to cell permeability issues or the inability of these compounds to fatally inhibit protein kinases in *T. brucei*.

For two reasons it may be interesting to investigate this PDK1-like protein kinase further; it may represent conservation of a key eukaryotic signalling pathway in an early branching eukaryote and may also represent a druggable target that has had many potential inhibitors synthesised by medicinal chemists attempting to target human PDK1.

5 DFK – A Negative Regulator of Bloodstream Form to Procyclic Form differentiation

5.1 Introduction

5.1.1 Trypanosome Differentiation

Trypanosoma brucei inhabits several very different environments during the course of its life cycle within various tissues of the mammalian host and the tsetse fly vector. It has to adapt to these environments and also ensure it does not kill the host before it can be transmitted to a new one. Early studies clearly noted that bloodstream form trypanosomes existed in differing morphological characteristics, slender and stumpy, with these strains being described as pleiomorphic (Vickerman, 1965; Steverding, 2008). Eventually, many laboratory cell lines lost this feature and became monomorphic - incapable of forming stumpy forms. This was due to the method of continuous mechanical passage of the infection, which selects for parasites with increased replication dynamics (Böhringer and Hecker, 1974) (G. Cross, <http://tryps.rockefeller.edu/>). The stumpy form of the parasite has been shown to be central to the transmission cycle of *T. brucei* as it is the form that can most efficiently differentiate from bloodstream form to the procyclic form in the tsetse fly.

The proliferative slender bloodstream form must pre-adapt in order to regulate its density and achieve transmission effectively, by forming stumpy bloodstream forms. This process occurs in a density-dependent fashion in response to an undefined small-molecule(s) termed stumpy induction factor (SIF). Slender to stumpy differentiation triggers changes which prepare the parasite for uptake by a tsetse fly and also puts the cell into a quiescent state (G_0/G_1 arrest) (Reuner et al., 1997; Vassella et al., 1997). This proposed role for the cell cycle ensures only certain cells commit to becoming stumpy and then allows for rapid and synchronous differentiation to PCF forms (Ziegelbauer et al., 1990; Matthews and Gull, 1994).

This density-dependent restriction of proliferation ensures control over parasite concentrations within the bloodstream. If this is not controlled, as is seen in animal infections with monomorphic laboratory strains, the proliferation will

increase in an unabated manner that proves fatal to the host. Once cells enter the stumpy form there is no return to proliferation unless the parasites undergo differentiation to procyclic form. The SIF trigger to form stumpy from slender cells has yet to be identified but was thought to activate cyclic nucleotide signalling pathways as it could be induced by cAMP analogues. However, it is now thought to be probably mediated by breakdown products of cyclic nucleotides, as cell permeable cyclic AMP analogues can be used to drive stumpy formation (Vassella et al., 1997; Laxman et al., 2006).

The pre-adaptations to transmission that occur consist of mitochondrial remodelling and elaboration to allow for aerobic respiration instead of glycosomal glycolysis, due to the main energy source of the procyclic forms being proline rather than glucose. Expression of VSG is dysregulated, potentially as this is not required for survival after transmission (Amiguet-Vercher et al., 2004). Uptake of haptoglobin-haemoglobin is reduced due to haptoglobin-haemoglobin receptor (HpHbR) expression being downregulated, though endocytosis levels of several other ligands remain high. The lysosome is also repositioned anterior of the nucleus (Vanhollebeke et al., 2010).

The surface of the trypanosome becomes further modified to contain PAD proteins PAD 1 and PAD 2. PAD 1 and PAD 2 are sited in an array of 8 PAD genes. PAD 1 is expressed only in stumpy forms with PAD 2 expressed in stumpy forms and procyclic forms, and thus PAD 1 expression provides a molecular cytological marker of stumpy forms. The genes encode proteins possessing an estimated 14 transmembrane domains and have homology to carboxylate transporters; when expressed in *Xenopus* oocytes they function as citrate transporters, supporting this notion. Depletion of PAD 1 by RNAi blocked the differentiation of BSF to PCF by addition of citrate:cis-aconitate (CCA) (Dean et al., 2009). CCA is probably the *in vivo* trigger for BSF to PCF differentiation; it can be used to differentiate BSF to PCF in culture and the efficacy of this increased by exposing parasites to cold shock (of 20°C, the approximate temperature experienced by dusk feeding tsetse flies) (Czichos et al., 1986; Engstler and Boshart, 2004). Cold shock increases the sensitivity to CCA of BSF trypanosomes to physiologically relevant levels, and also PAD 2 levels are upregulated by cold shock (Engstler and Boshart, 2004; Dean et al., 2009).

The differentiation to procyclic forms results in major morphological changes. The cells have a visibly different morphology and motility to BSF, the mitochondrion is enlarged and branched, and the positioning of the kinetoplast moves closer towards the nucleus (away from the posterior of the cell). Cell surface glycoproteins are changed, with a family of procyclins replacing the VSG coat. Procyclins play a role in protecting the parasites from peptidases in the tsetse midgut (Roditi et al., 1989). Two major forms of procyclin exist and are termed EP or GPEET, depending on the amino acid repeats they contain (Glu-Pro or Gly-Pro-Glu-Glu-Thr respectively) and multiple forms of these major types exist (Mowatt and Clayton, 1987; Mowatt et al., 1989). In fly infections the expression of procyclins is, to some extent, sequential with GPEET being expressed in early infections then transitioning to the various EP procyclin in late procyclic forms (with a concomitant reduction in GPEET); procyclins are subject to some proteolytic degradation but have a protease-resistant C-terminal domain (Acosta-Serrano et al., 2001). In culture the ratio of EP to GPEET can vary between different cell lines (Bütikofer et al., 1997). The levels of expression of procyclins (particularly GPEET) can further be regulated by the activity of mitochondrial enzymes (Vassella et al., 2004).

Several studies have attempted to quantify the global changes in mRNA levels that occur when the BSF parasites differentiate to stumpy and procyclic forms. Microarray analysis of a portion of the trypanosome genome (796 genes) suggested that around 6% of this cohort was differentially regulated between BSF and PCF. Though this study was rapidly superseded by whole transcriptome analyses, it did detect changes in transcript levels of some protein kinases (Koumandou et al., 2008). Three studies used microarray analysis to try and quantify the changes between BSF and PCF and during the differentiation from slender to stumpy to procyclic. Various cell lines were used in each experiment, differentiation was conducted in different manners and variation between each data set was significant. However, all studies at least agreed there was significant remodelling of the transcriptome during differentiation with 14-20% of genes being differentially regulated (Jensen et al., 2009; Kabani et al., 2009; Queiroz et al., 2009). Of relevance to this study, the previous works all detected differences in protein kinase and phosphatase transcripts. Characterisation by RNA-Seq of the BSF and PCF transcriptome showed around 6% of transcripts were

differentially regulated two-fold or more, and was broadly in agreement with other studies (Siegel et al., 2010). Also broadly in agreement with this is the comparison of BSF and PCF proteomes by stable isotope labelling in culture (SILAC) followed by sensitive LC-MS/MS, showing 10% of the proteome being differentially regulated, but at much greater fold change (up to 5-fold)(Urbaniak et al., 2012a). Both triggering and effecting the differentiation process in trypanosomes requires signalling, and reversible phosphorylation has been implicated and well studied in this process.

5.1.2 Reversible phosphorylation during trypanosome differentiation

Differential phosphorylation of tyrosine and differential tyrosine kinase activity has been demonstrated in the various developmental stages of trypanosomes, though without assignation of functional relevance (Parsons et al., 1991, 1993). The first functional implication of a protein kinase in differentiation came from analysis of a protein kinase in the AGC family. Initially investigated for its unusual domain architecture - a PX domain followed by a FYVE zinc finger then the protein kinase domain. When the gene was knocked out of monomorphic BSF trypanosomes (MiTat 1.4) no growth defect was evident. When it was knocked out of pleomorphic AnTat 1.1 BSF cells they began to develop growth retardation after around 50 hours in culture (after obtaining cells from infected mice and pre-culturing them in methylcellulose containing medium). Growth then began to plateau after 4 days. It was observed that the cells were differentiating to stumpy forms, even though cell density was being controlled to prevent this. Differentiation was confirmed by assaying for NAD diaphorase activity (detecting the increase in mitochondrial activity), with about 90% of the cells being diaphorase positive and in the G₁/G₀ phase of the cell cycle. The knockout parasites exhibited equal virulence to unmodified parasites when used to infect both immunocompetent and immunosuppressed mice and could also be transmitted through tsetse flies which could then re-infect mice. The increased sensitivity to parasite density indicates that ZFK has a dampening effect on the activity of SIF (Vassella et al., 2001).

The second protein kinase to be implicated in the slender to stumpy differentiation was TbMAPK5 (Domenicali Pfister et al., 2006a). MAPKs or

mitogen activated protein kinases are CMGC family kinases and can be further subdivided by the central amino acid in the activation loop TXY motif. TbMAPK5 has a TDY motif, distinguishing itself from MAPKs in higher eukaryotes; phylogenetic analysis has also placed it outside of known MAPK classes. When knocked out of monomorphic BSF cells no growth defect was observed (as for ZFK) but when *in vitro* differentiation to PCF was conducted the knockout had a 50% lower efficiency. When the knockout was conducted in pleiomorphic BSF and the cells used to infect mice a growth defect was observed. Peak parasitaemias were lower in the knockout cell lines, though waves of parasitaemia were still observed. In order to control for the effect of the mouse immune response (in case MAPK5 knockouts were more susceptible to it) the lines were used to infect immunosuppressed mice. Compared with the wild type the knockout lines did not increase their maximal parasitaemia, suggesting they were more sensitive to cell density and that this was limiting their growth. Parasites containing an add back partially restored their virulence, but when an add back TbMAPK5 with the TDY motif mutated to ADF (making it inactivatable, and therefore dead or much reduced in kinase activity) was added, the phenotype was much more pronounced. Microscopic analysis showed the cells were forming morphologically stumpy forms at much lower densities (and was confirmed by diaphorase staining). This was replicated in culture by allowing cells to grow in unconditioned medium or media mixed to contain various percentages of conditioned medium, revealing that TbMAPK5 knockout cells showed a much higher sensitivity to SIF.

The result of these two studies demonstrated a role for ZFK and MAPK5 in controlling the responsiveness of trypanosomes to SIF by repressing them from differentiating to slender from stumpy cells at low densities. Recently, another protein kinase was implicated in this process (Barquilla et al., 2012). TbTORC4, a complex containing the TbTOR4 kinase (a PIKK (phosphatidyl inositol 3' kinase-related kinases, atypical protein kinase group)) was found to negatively regulate stumpy formation in monomorphic trypanosomes. Transient or continuous RNAi of TbTOR4 led to irreversible differentiation to a G0/G1 arrested state. TbTOR4 RNAi induced cells became diaphorase positive, indicating increased mitochondrial oxidative capacity, became resistant to mild acid and developed a shortening of the flagellum with a rounded posterior, consistent with stumpy

form trypanosomes. Transcript profiling was conducted for a set of 10 genes known to be selectively upregulated in stumpy forms and for the majority of these the change in TbTor4 RNAi cells vs. parental cells was the same as stumpy vs. slender cells, suggesting the cells had developed a transcriptome similar to stumpy forms. RNA polymerase 1 was shown to delocalise from the expression site body to all over the nucleus, again consistent with a stumpy phenotype.

The transcripts upregulated in TbTOR4 RNAi lines included PAD 1 and PAD 2, which corresponded to an increase in the expression of PAD1 protein by western blot. This protein was presumably being trafficked to the cell membrane as the RNAi lines became hypersensitive to physiologically relevant levels of *cis*-aconitate, with approximately 20% of cells gaining EP procyclin after exposure to 0.01 mM or 0.1 mM *cis*-aconitate. Taken together, this suggests that TbTOR4 depletion is required for pre-adapting the parasites to allow effective transmission. The study sought to assess what signalling pathways might be regulating TbTOR4 expression by treating the cells with hydrolysable and non-hydrolyzable cAMP analogues and AMP analogues. AMP and hydrolysable cAMP analogues induced the downregulation of TbTOR4 expression, though non-hydrolysable analogues of cAMP did not. The authors conclude that AMP levels, or specifically a high ratio of AMP to ATP (thus indicating reduced cellular energy), inactivate TbTOR4 to trigger differentiation of slender to stumpy forms. The authors also suggest that TbMAPK5 and ZFK also participate in the same complex signalling pathway, and that as ZFK is an AGC kinase (which is a well-known TOR substrates in other organisms) it may be acting as a downstream effector of TbTOR4. Research in other organisms has linked TOR kinase functions to nutrient sensing and TbTOR4 may represent a trypanosome specific variation on this process (Howell and Manning, 2011). This may be supported by the fact that inhibition of glucose transport into BSF cells, and thus nutrient stressing them, can induce a partial differentiation to PCF forms (Haanstra et al., 2011).

The next stage of differentiation, from stumpy to PCF, is a process that also had a MAP kinases implicated in its progression (Müller et al., 2002a). When TbMAPK2 was knocked out of monomorphic bloodstream form parasites no difference in the rate of growth or morphology was observed. When the BSF cells were induced to differentiate to PCF cells by the addition of *cis*-aconitate they appeared to do so (by morphology) but subsequently became growth arrested.

Specificity of the phenotype was confirmed by TbMAPK2 add back at an exogenous locus. Examination of the phenotype in more detail showed that the cells were experiencing a delay in the differentiation process. This was assessed by the expression of PCF markers such as EP procyclin (early) and CAP5.5 (a late procyclic marker), which were both delayed to the same extent, suggesting differentiation proceeded at a normal pace but cells were delayed in their licensing for differentiation. After differentiation cells were shown to go through approximately one cell division but then became growth arrested thereafter, and also exhibited abnormal nucleus:kinetoplast conformations. The re-expression of TbMAPK2 with a non-activatable AEF (instead of TEY) motif in the activation loop generated the same phenotype, suggesting that TbMAPK2 is also dependent on upstream MAPKKs for activation. Despite many attempts the gene could not be knocked out of established PCF cell lines, further indicating that it plays an essential role in the proliferation of PCF parasites.

As protein kinase activity is clearly important to the differentiation process it is reasonable to assume that protein phosphatase (enzymes that dephosphorylate a substrate) activity is also involved at some stage. *T. brucei* possesses 78 proteins containing a protein phosphatase domain (Brenchley et al., 2007). Whilst searching for genes likely to be involved in stumpy BSF differentiation competence Szöör *et al.* (2006) identified a gene with similarity to a mammalian protein tyrosine phosphatase (PTP) that they investigated further (Szöör et al., 2006). Named TbPTP1, it was found to be upregulated in stumpy forms at a transcript level, but that protein levels remained constant throughout slender, stumpy and PCF stages. It was found to predominantly associate with the cytoskeletal fraction of differentially detergent-fractionated BSF cells. Analysis of the amino acid sequence showed that TbPTP1 contained 9 out of 10 PTP-specific conserved residues and motifs compared with human PTP, and possessed a trypanosome specific motif replacing the missing conserved motif. Other trypanosome specific motifs also led the authors to place TbPTP1 in a trypanosome specific PTP family. TbPTP1 was shown to be a functional tyrosine-specific protein phosphatase *in vitro* and did not exhibit activity against Ser/Thr-phosphorylated substrates, phospholipids, nucleotides or inorganic phosphocompounds. The recombinant enzyme could be inhibited with tyrosine phosphatase specific inhibitors. Most interestingly was that when TbPTP1 was

depleted by RNA interference between 2-12% of cells, of a monomorphic stock, spontaneously differentiated to PCF cells at 37 °C in HMI-9 medium. The cells stained positively with anti-EP, anti-GPEET and anti-CAP5.5 antibodies, and exhibited PCF characteristic repositioning of the kinetoplast. Although the RNAi phenotype was unstable (due to the use of pZJM system), an inhibitor of mammalian PTP1B, BZ3, could induce the same phenotype as TbPTP1 RNAi when applied at 150 µM. BZ3 was also shown to inhibit the recombinant TbPTP1. The specificity of BZ3 for TbPTP1 was further demonstrated by inducibly expressing a dominant-negative, substrate trapping mutant (D199A) and treating the cells with BZ3, compared with the uninduced control the substrate trapping mutant and BZ3 treatment led to an increase in the proportion of cells differentiating to PCF. This demonstrated that BZ3 was not inhibiting a downstream target of TbPTP1.

The authors then used BZ3 to investigate the role of TbPTP1 in the differentiation in pleiomorphic cells. When applied to slender parasites extracted from infected mice, BZ3 induced a very similar differentiation phenotype as seen when applied to monomorphic BSF cells, i.e. 7% differentiation after 24 h. When BZ3 was applied to stumpy enriched populations extracted from mice the phenotype was strikingly similar to that induced by the addition of *cis*-aconitate. Both these experiments were conducted at 37 °C in HMI-9, and cells expressed EP procyclin, CAP5.5 and repositioned the kinetoplast to procyclic type morphology. The authors account for the differentiation of monomorphic and slender pleiomorphic cells when TbPTP1 is inhibited by inferring there is a population of Stumpy* cells that respond to TbPTP1 inhibition. Stumpy* cells are considered slender in morphology but have the functionality of a stumpy in terms of their ability to differentiate (Tasker et al., 2000; Matthews et al., 2004).

The ability to generate substrate trapping mutants of TbPTP1 allowed further investigation of its biological role (Szöör et al., 2010). By conducting a pulldown using resin-chelated TbPTP1-D199A a protein of 39 kDa was found to interact. This was confirmed using reciprocal pulldown and the protein was named TbPIP39 (TbPTP1-interacting protein, 39 kDa). The authors identified a potential phosphorylation site (Tyr 278) and used FGR kinase (human Gardner-Rashid feline

sarcoma viral (v-fgr) oncogene homologue protein kinase) to phosphorylate this residue (confirmed by generating a T278F mutant). Phosphorylated TbPIP39 could then be dephosphorylated by TbPTP1 but not by an inactive mutant of TbPTP1 *in vitro*, indicating that it was also a potential substrate of TbPTP1 as well as an interacting partner. This tyrosine residue was also deduced to be important for the activity of TbPIP39, since while phosphorylated the protein exhibited greater phosphatase activity than when not. When the two phosphatases were combined to dephosphorylate a third substrate a synergistic effect was observed and determined to be due to TbPIP39 increasing the activity of TbPTP1. This activity could be abolished by the addition of CCA. TbPIP39 was shown to localise to glycosomes and that it was only expressed as a protein in stumpy or PCF cells. By testing cell lysates with an antibody against phosphorylated TbPIP39 Y278 it was shown that under *cis*-aconitate or BZ3 treatment phosphorylation of this residue was enhanced. Expression of TbPIP39 was also upregulated by cold shock. These factors led the authors to predict that TbPIP39 was involved in regulating the stumpy to PCF differentiation process in some manner. To test the hypothesis of TbPIP39 being involved in the stumpy-PCF differentiation it was ablated by RNAi and differentiation induced by addition of *cis*-aconitate. At physiologically relevant levels of CCA the differentiation process was reduced in RNAi induced cultures, implicating TbPIP39 as being an effector of stumpy to PCF differentiation. Re-expressing a wild-type but recoded TbPIP39 allele (which was therefore refractory to RNAi) restored the ability to differentiate at 0.1 μM *cis*-aconitate. Expression of mutants that were incapable of entering the glycosomes could not compensate the differentiation defect, nor could expression of a catalytically dead mutant. Taken together the results of Szöör *et al.* (2006, 2011) demonstrate that a phosphatase cascade is responsible for triggering effective BSF to PCF differentiation.

The differentiation process could provide a novel mechanism of therapeutically targeting the parasite. Using an inhibitor to induce differentiation to stumpy or PCF cells in the bloodstream of the host would eventually prove fatal to the parasite. If parasites could be efficiently stimulated to enter stumpy G_0/G_1 arrest, using an inhibitor of TbMAPK5 for example, they would eventually perish due to their inability to re-enter proliferation. Aberrant differentiation to PCF

forms would also be fatal to the parasite, as the expression of procyclin could make them susceptible to complement mediated lysis and also provides a static immunological target (due to loss of VSG expression and switching) (Ferrante and Allison, 1983). However, this would necessitate any inhibitor to cause 100% switching of parasites to a stumpy or procyclic form to be effective.

5.1.3 Research Aims

As reversible protein phosphorylation is clearly implicated in trypanosome differentiation it was evident that the pTL library provided a resource for investigating this process further. The 2T1 strain of trypanosome is considered to be monomorphic so it was reasoned that stumpy formation would be difficult to investigate. However, the 2T1 cell line can differentiate asynchronously to PCF using *cis*-aconitate stimulation. Two possible roles can be envisioned for protein kinases involved in regulating BSF-PCF differentiation; negative and positive regulation of differentiation. The first scenario is that RNAi of a protein kinase results in the spontaneous induction of BSF to PCF differentiation, akin to the TbPTP1 RNAi phenotype. The second is that RNAi of a protein kinase will not generate an obvious phenotype in BSF cells but when they are induced with *cis*-aconitate are not able to progress to PCF cells. The first scenario was the easiest to test during our initial RNAi screen, and was carried out by observing cell morphology during RNAi, and/or testing for procyclin expression. The second scenario could not be tested due to time and labour constraints.

Following identification of any such protein kinase involved in regulation of BSF-PCF differentiation studies would be undertaken to attempt to determine the function of the gene including characterisation of the phenotype, localisation of the protein, attempts at functional genetics and analysis of recombinant protein.

5.2 Results

5.2.1 Identification and Features of DFK

Whilst undertaking the kinome-wide RNAi screen an RNAi cell line, targeting the gene Tb11.01.5650, was observed to exhibit a morphological defect under tetracycline induction where the BSF parasites exhibited an elongated shape and

began to move in an uncharacteristic manner. Because of the interesting phenotype this cell line was investigated further. Based on the phenotype of the RNAi mutant this protein kinase was eventually termed DFK for Differentiation Kinase, the reasons for this are described below. This phenotype was not observed in any of the other protein kinase RNAi lines.

Tb11.01.5650 was detected in the bioinformatic classification of the kinome as a STE11 like kinase (Parsons et al., 2005b). The CDS of the gene is 3588 nt in length and this encodes a 1195 amino acid peptide - a protein with a predicted mass of 131 kDa. The DFK protein is predicted to have a kinase domain in the very C-terminal region with most of the protein being N-terminal to the kinase domain (Figure 5-1). This region is predicted to contain multiple transmembrane domains (Table 5-1). However, different prediction algorithms gave differing numbers of TM domains, and predicted them in different positions of the sequence (Figure 5-2). DFK has been detected in the flagellar membrane proteome, suggesting it is present in the flagellar or cell membrane (Oberholzer et al., 2011). Other domains were detected using PFAM analysis. These were a HAMP domain and a TPR2 domain - which had significant E-Values and are also present in the *T. cruzi* orthologue (TcCLB.511727.210) - which was originally identified by Parsons *et al.* 2005. HAMP stands for “Present in Histidine Kinases, Adenylate Cyclases, Methyl Accepting Proteins and Phosphatases” and is a linker domain that crosses the plasma membrane (Parkinson, 2010). TPR domains (Tertatricopeptide-like Repeats) are found in a huge diversity of proteins and are thought to mediate protein-protein interactions (Allan and Ratajczak, 2011; Zeytuni and Zarivach, 2012). To our knowledge the protein has not been detected as being post-translationally modified (TriTrypDB).

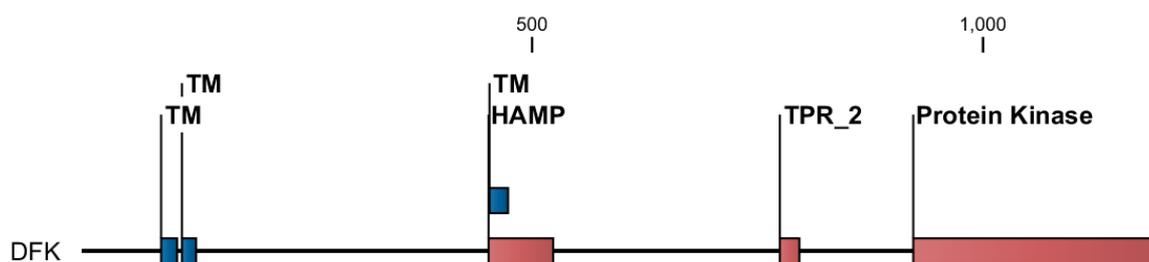


Figure 5-1: Schematic representation of the DFK protein: TM domains are highlighted corresponding to those predicted by the TMHMM algorithm. HAMP, TPR and protein kinase domain determined by PFAM domain search. Scale above indicates amino acid position.

Table 5-1: List of the transmembrane prediction algorithms used to assess the potential TM domains in DFK.

Algorithm	Reference	Number of TM Domains
TMHMM	http://www.cbs.dtu.dk/services/TMHMM/	3
HMMTOP	www.enzim.hu/hmmtop/	3
TM-PRED	www.ch.embnet.org/software/TMPRED_form.html	5
SOSUI	http://bp.nuap.nagoya-u.ac.jp/sosui/	2
PRED-TMR	http://athina.biol.uoa.gr/PRED-TMR/	4

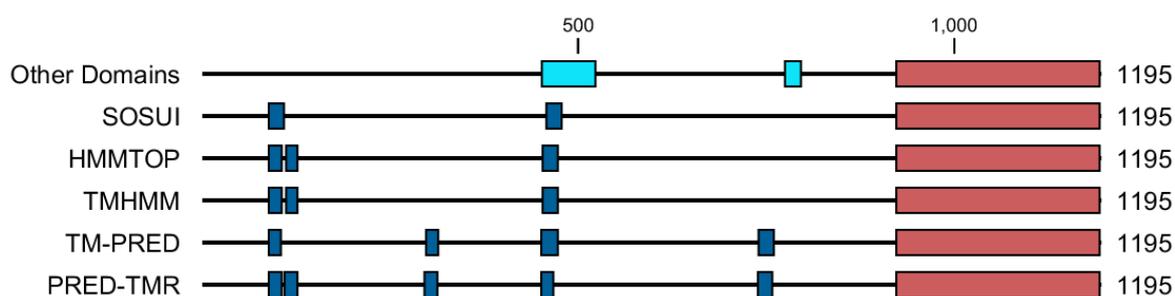


Figure 5-2: Diagram of the position of predicted TM domains of DFK from five different prediction algorithms (blue), which are all in the region N-terminal to the protein kinase domain (red). The top sequence shows other domains detected in PFAM analysis using CLC genomics – the left cyan bar represents a HAMP domain and the right a TPR repeat.

5.2.2 RNAi Results

RNAi ablation of DFK mRNA in BSF trypanosomes in HMI-11 at 37°C did not generate a growth phenotype detectable by the alamar blue assay. However, when viewed under a microscope at 72 hours post-induction the morphology of an obvious proportion of the cells was clearly not consistent with BSF trypanosomes, with the cells appearing to look and move like procyclic form trypanosomes. Using two bioreplicates (cell lines STL201/202) a cumulative growth curve was performed over several days, showing that the cells developed a slow-growth phenotype by 72 h that should have been detectable by the alamar blue screen. The growth defect became apparent at 48 h post-induction and this became more severe over the course of a 6 day experiment (Figure 5-3). The morphological phenotype was termed PCF-like prior to further confirmation. DFK RNAi cell lines were used to infect mice and RNAi was induced in the test mice by the addition of doxycycline to drinking water. Due to the weakness of the growth defect, RNAi induction was performed at the same time as infection in order to increase the chances of detecting the weak phenotype. Parasitaemia was unaffected in the test mice despite multiple attempts and varying the level

of inoculum and timing of RNAi induction (Figure 5-3b). Targeting DFK by RNAi in culture caused a 60 % depletion of the mRNA after 24 h (as determined by qPCR of DFK mRNA levels) (Figure 3-7). Western blotting against a tagged allele showed an almost total depletion of the tagged allele by 24 h (Figure 5-7), showing that the RNAi of DFK was effective.

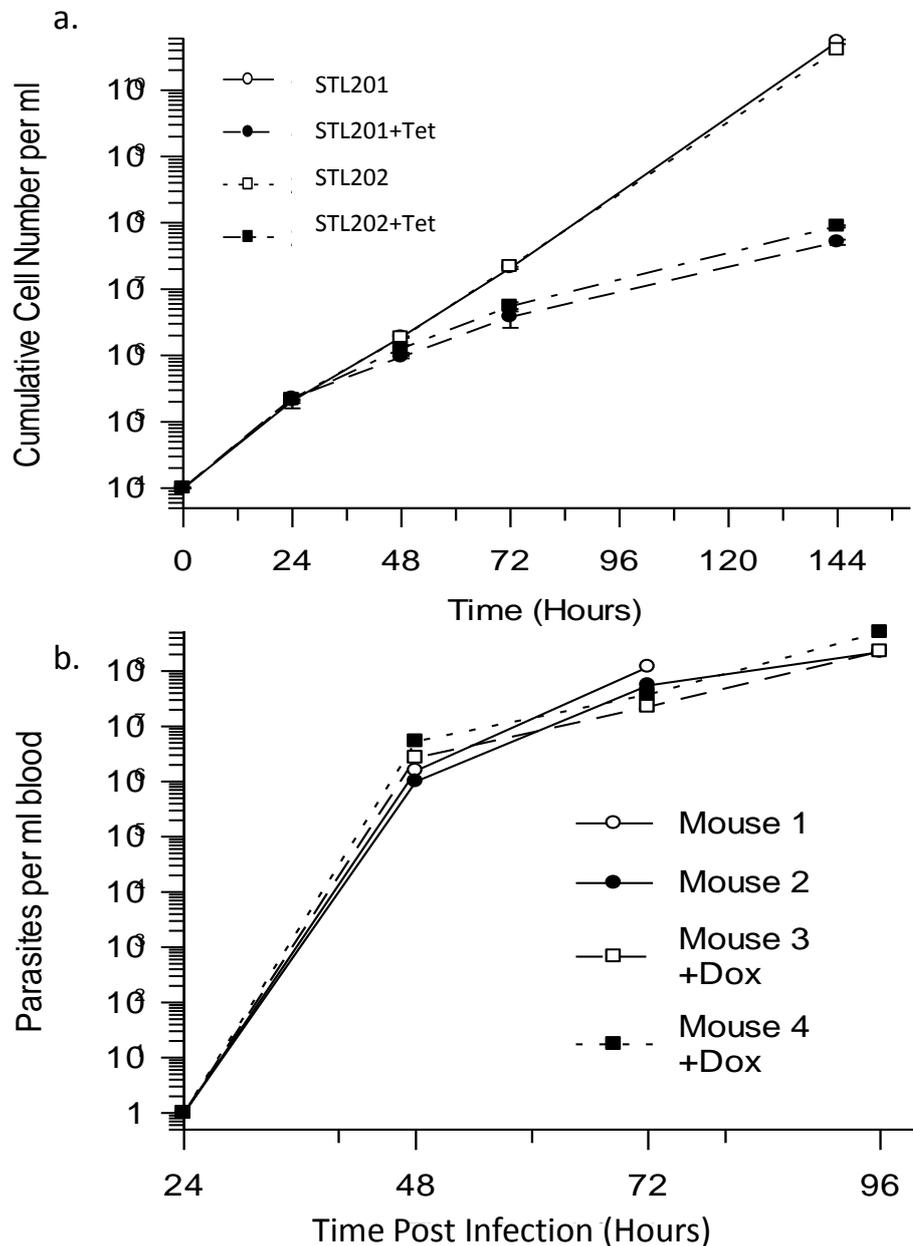


Figure 5-3: Growth curves of DFK RNAi cell lines (plus/minus induction). a. In culture RNAi using cell lines STL201/STL202 (two bioreplicates) with mean values plotted and 1 standard deviation error bars (3 technical replicates). b. Parasitaemia of infected mice using STL201, two mice (3 and 4) were given Doxycycline laced water to induce RNAi as soon as they were infected. The limit of detection of parasitaemia was 10⁴ parasites per ml of blood.

5.2.3 Evidence of Differentiation

As it appeared that the ablation of DFK RNAi was inducing a PCF-like morphology the expression of procyclic-specific markers was tested. Procyclic cells express EP procyclin so the DFK RNAi cell lines were tested for expression of this cell surface glycoprotein using a monoclonal anti-EP procyclin antibody conjugated to FITC. This allowed simple labelling protocols for both FACS (fluorescence activated cell sorting) and IFA (immunofluorescence analysis) applications.

In order to assess whether DFK RNAi cell lines expressed EP procyclin and thus to quantify the proportion of cells differentiating to PCF-like cells in the RNAi a FACS assay was performed. Cells were fixed, washed and were labelled with FITC-conjugated anti-EP procyclin before being assayed by flow cytometry using a FACSCalibur system. Control samples consisting of BSF 2T1 cells, PCF 927 cells and a mix of BSF and PCF parasites were used to get baseline data on a negative control, positive control and mixed sample respectively. The curves generated by these controls were used to set bifurcation gates to define the EP procyclin positive and negative parasites in the test samples (Figure 5-4). Samples were taken of cell lines in which DFK RNAi had been induced for 24 h, 48 h and 72 h timepoints and assayed (Figure 5-5). The results showed that 0.18% of events were EP procyclin positive in the 2T1 parental cell line, which indicates a level of background labelling in the assay. This increased to 0.35% in the uninduced DFK RNAi lines. In the induced DFK RNAi samples at 24 h, 1.85% of cells were EP procyclin positive, at 48h 5.04% were positive and at 72 h 16.7% were positive. For each sample 100,000 events were collected, gated to exclude debris and separated by bifurcation gate into EP positive and negative populations. The counts were evaluated by Chi square analysis for trend using Prism Graphpad which determined the increase in EP positive cells to be highly significant.

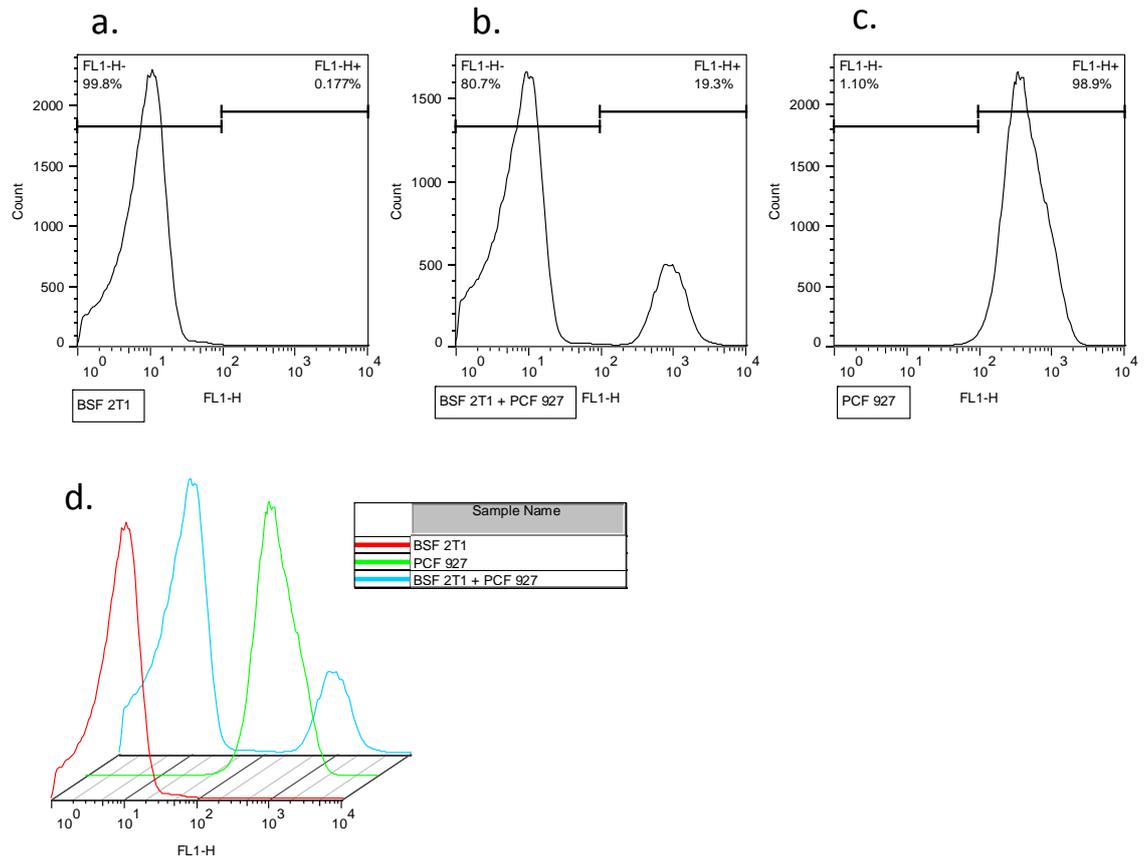


Figure 5-4: FACS curves for control samples showing the fluorescence of BSF (a.), PCF (c.) and mixed cells (b.) when labelled with a FITC-conjugated anti-EP procyclin monoclonal antibody, results are combined in panel d. The mixed BSF/PCF sample (b.) was used to define the bifurcation gate to separate PCF from BSF in further experiments. N=100,000 cells per experiment.

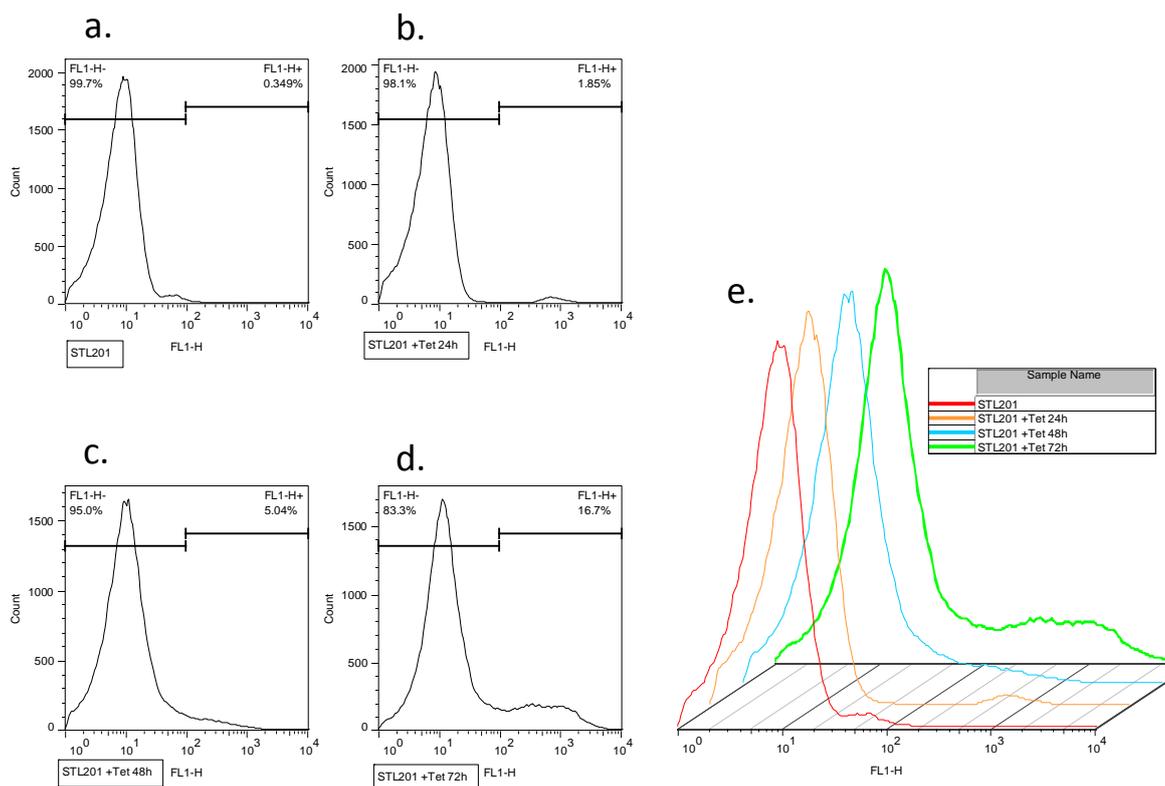


Figure 5-5: FACS curves for DFK RNAi (Cell line STL201) samples at 37°C in HMI-11 showing the fluorescence of RNAi induced cells when labelled with a FITC-conjugated anti-EP procyclin monoclonal antibody. This marker was used to determine differentiation to PCF. Uninduced control a. compared to induced samples at 24 hour intervals until 72 h (b.-d.) with panel d showing 17% PCF cells in the culture. Combined results are shown in panel e. Bifurcation gate set from Figure 5-4b. N=100,000 cells per experiment.

Cells that had been analysed by FACS were further stained with DAPI and observed under a fluorescence microscope to assess the morphology of the differentiated cells. It was observed that a number of the EP procyclin positive cells were undergoing kinetoplast repositioning to exhibit a procyclic morphology (Figure 5-6a/b). Other cells were clearly abnormal in their morphology and nucleus/kinetoplast configuration though no attempts to quantify a specific defect have yet been made. The defect may be specific to DFK RNAi or may be due to 2T1 cells not having the ability to proliferate effectively as PCF cells. Images collected using an IN-Cell Analyser 2 supported the FACS data as approximately 20% of cells were EP procyclin positive in a test of cultures where RNAi was induced for 96 h (E. Kalkman, Scottish Bioscreening Facility). Repositioning of the kinetoplast was also observed in these samples as well as undefined morphological and nucleus-kinetoplast positioning defects (Figure 5-6c). Although no quantitative analysis of this data has yet been performed it should be possible to write an algorithm using IN-Cell Explorer software to measure nucleo-kinetoplast distances in the induced and uninduced samples.

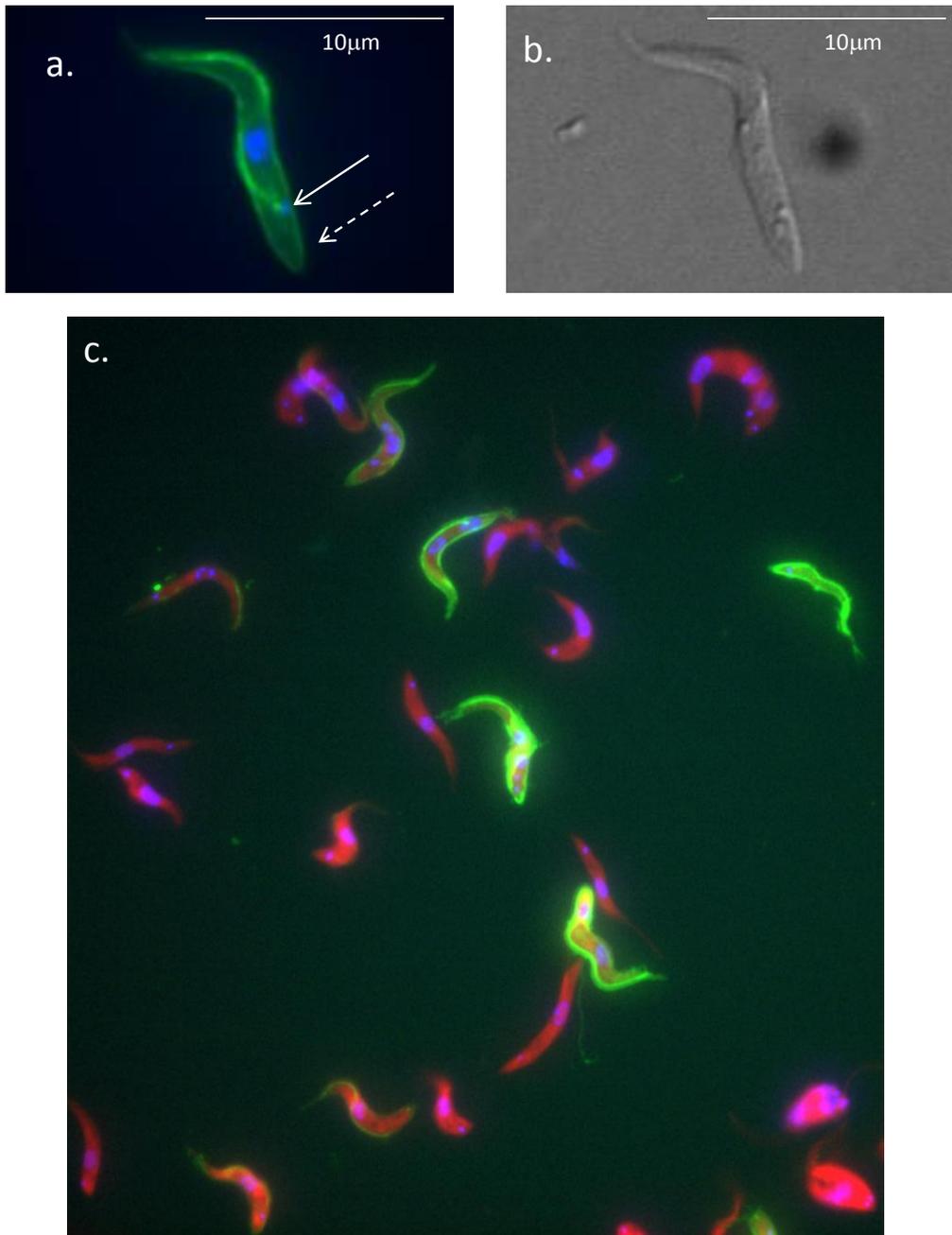


Figure 5-6: a. An example of an apparently morphologically correct PCF after differentiation from BSF induced by RNAi of DFK for 72 h (a. IFA, b. DIC) Green: FITC conjugated anti-EP procyclin, Blue: DAPI. Panel c. Shows a representative, merged image from the IN-Cell 2000 of DFK RNAi after 96 h of induction. Green: FITC conjugated anti-EP procyclin, Blue: DAPI, Red: Cellmask Red cytoplasmic stain. In panel a. solid arrow denotes kinetoplast repositioning from the typical posterior location in a BSF parasite (approximate location indicated with dashed arrow).

In order to assess if the differentiation process was occurring when RNAi of DFK was induced in a mouse infection, terminal bleeds were taken from control and test mice, and the buffy coat separated from whole blood. This was washed with PBS and labelled with the anti-EP procyclin antibody prior to analysis by IFA and

FACS but no evidence of differentiated PCF parasites could be detected these isolates.

5.2.4 Localisation of DFK

5.2.4.1 Epitope Tagging and Crude Cell Fractionation

An allele of DFK was endogenously tagged with a C-terminal 12Myc epitope using the pNAT BSD x12Myc (pGL2178) vector (Alsford and Horn, 2008). The final 3' 1kb of the CDS (minus the stop codon) of the DFK gene was cloned into this vector between the HindIII and XbaI sites of the vector (pGL2203). The plasmid was linearised with an insert unique PstI site and transfected. This tagging technique inserts the 3' of the CDS fused to the 12Myc epitope tag followed by a stop codon into the endogenous DFK locus. The 3' UTR is replaced with an actin UTR in this strategy. The plasmid was transfected into cells carrying the DFK RNAi cassette (pTL129, STL403). Eight blasticidin resistant clones were screened for expression of Myc by western blot. Three were found to express high levels of a Myc-tagged protein close to the predicted size of DFK-12Myc and two (clones 1/2) were selected for further analysis (designated STL2378/2379) (Figure 5-7a). To provide further evidence that it was DFK that had been tagged and that DFK RNAi was effective an RNAi induction was performed. After 24 h induction, the signal on an anti-Myc western blot was almost totally ablated from both cell lines (Figure 5-7b).

As DFK was observed in the flagellar membrane proteome (Oberholzer et al., 2011) and has multiple transmembrane domains a crude cell fractionation was performed to give a low resolution indication of where DFK localises (Figure 5-7c). DFK12Myc expressing BSF parasites were lysed using a hypotonic TrisHCl buffer or a Triton X-100 containing buffer to differentially fractionate the cells. DFK-12Myc was detected in a soluble fraction of the cell following lysis in a Triton X-100 detergent-containing buffer and an insoluble pellet from a hypotonic cell lysis, suggesting a membrane location. It was not detected in fractions corresponding to the cytosol or cytoskeleton in either detergent or hypotonic lysis. The anti-Myc signal obtained on the western blot against the insoluble pellet from hypotonic lysis was dramatically reduced. To test for the presence of Oligopeptidase B (OPB), which should localise to the cytosolic

fractions, sheep anti-OPB serum was used as a control (Burleigh et al., 1997; Morty et al., 2001; Munday et al., 2011). The KMX-1 antibody was used to probe for β -tubulin, which should be found in the cytoskeletal fraction. β -tubulin was also detected in the cytoskeletal and cytosol containing fraction of the hypotonic lysis sample. This is at odds with previous work but does not affect the interpretation that DFK is a membrane associated protein. This effect may be due to fragmentation of the cytoskeleton by the sonication used during lysis, allowing some β -tubulin to enter solution. Attempts were made to use the Pierce M-PER kit, which is designed to purify integral membrane proteins, to purify DFK-12Myc. However, the volumes of reagents used are quite high and appear to require very large numbers of parasite cells to generate concentrated protein extracts. Optimisation of this procedure was not sufficiently advanced to obtain membrane extracts of sufficient concentration to provide a signal strong enough for western blotting experiments.

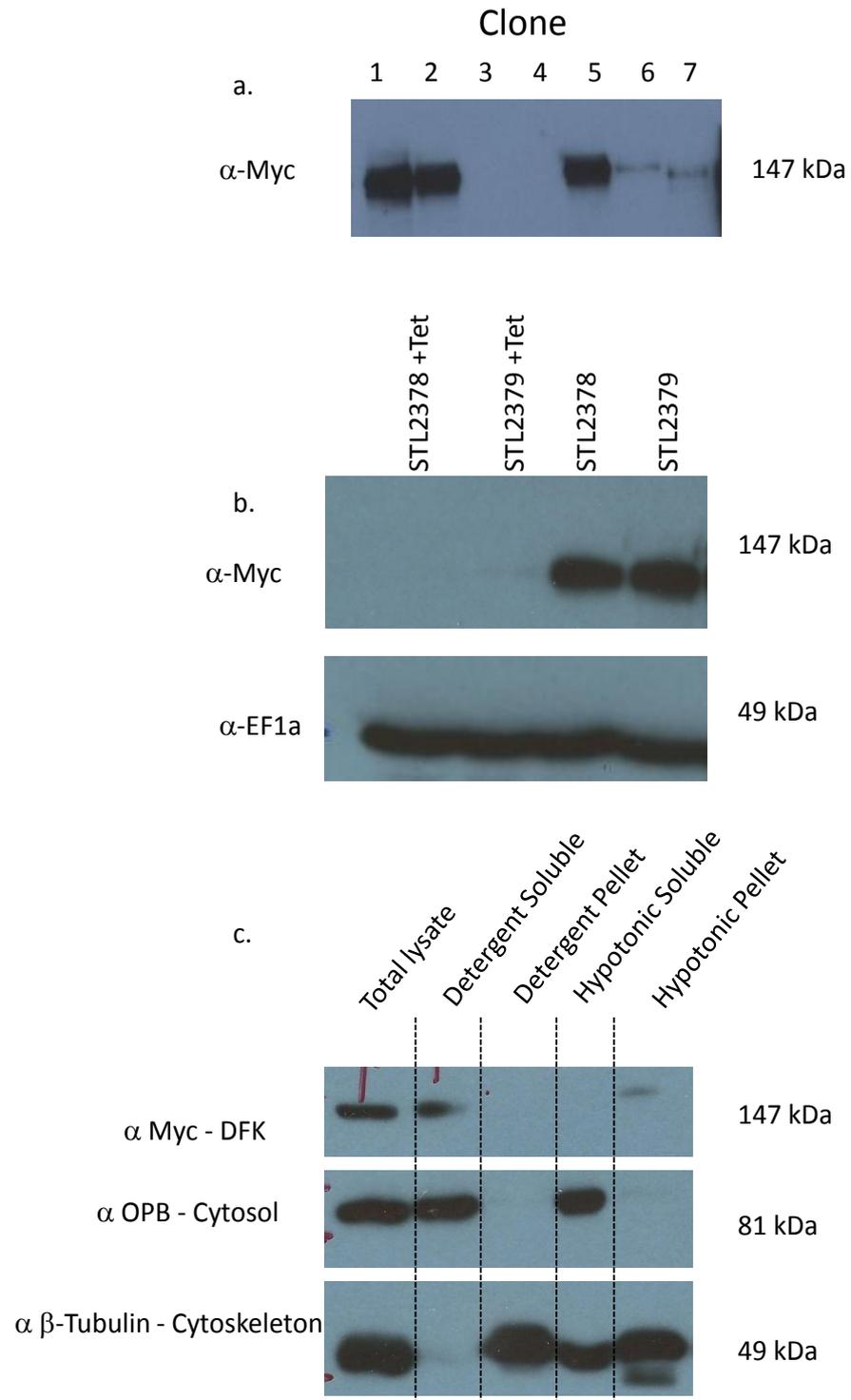


Figure 5-7: Panel a. Western blot to screen parasite clones transfected with pGL2203 to endogenously tag DFK with C-terminal 12Myc. Panel b. shows the results of a western blot against cell lysates from the selected clones STL 2378/2379 which express a DFK-12Myc fusion protein. EF1- α provides a loading control. Panel c. is the result of a western blot against crude sub-cellular fractionations of STL2378. OPB is used as a control for a cytosolic protein and β -tubulin a control for a cytoskeletal protein.

5.2.4.2 Immunofluorescence Analysis of DFK-12Myc Tagged Cell Lines

In order to gain further insights into the localisation of DFK, the DFK-12Myc cell lines were subjected to immunofluorescence analysis (IFA). Cells were dried onto multiwell slides coated in poly-L-lysine and fixed with 4% formaldehyde PBS v/v. Some samples were treated with 0.1% Triton X-100 PBS v/v to increase membrane permeabilisation for comparison. Samples were blocked with 1% BSA PBS w/v. Following labelling with mouse anti-Myc (4A6, Millipore), anti-mouse AlexaFluor-594 and DAPI, cells were observed under an epifluorescence microscope. Control samples were included to determine if the secondary antibody binding was specific, and that the primary antibody did not bind to cells not expressing the Myc epitope tag. Cells not exposed to detergent stained brightly and evenly over the whole cell membrane, including the flagellum. Those that had been permeabilised exhibited a weaker and more punctuate fluorescence pattern, occasionally with brighter staining in an area that could represent the flagellar pocket. All the cells in the sample appeared to be equally labelled. Cells that did not express DFK-12Myc did not label with the primary antibody. Cells not treated with the primary antibody did not bind the fluorescently labelled secondary antibody. These data are consistent with DFK exhibiting a cell membrane localisation.

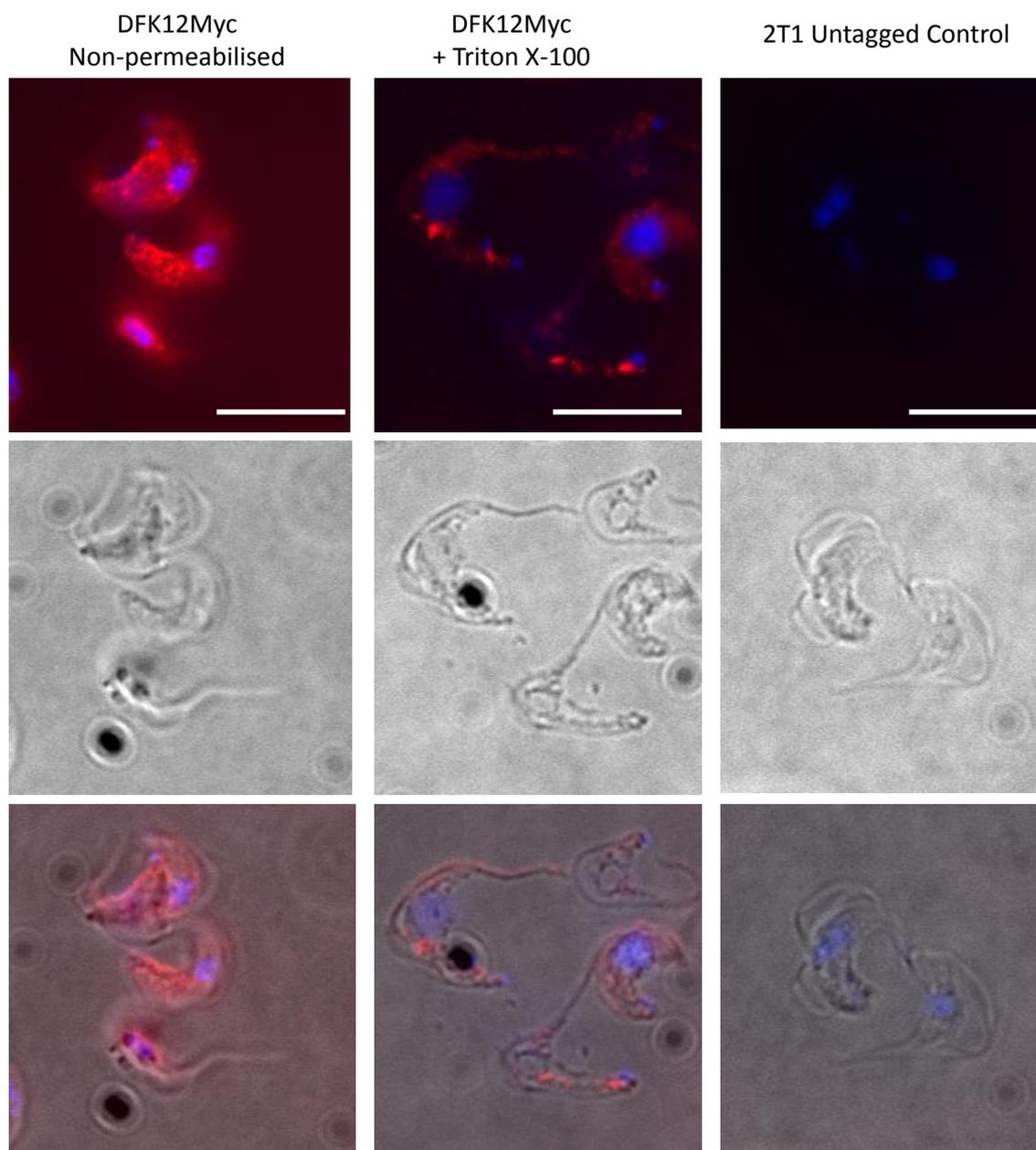


Figure 5-8: Immunofluorescence localisation of DFK-12Myc in cell line STL2379, using 2T1 cells as a negative control. Prior to immunolabelling cells were permeabilized using Triton X-100 or left intact before incubation with anti-Myc antibody and anti-mouse Alexa Fluor 594 conjugate visualised using appropriate filters (red). DNA was stained using DAPI (blue). Reference images were taken using DIC, allowing fluorescence and DIC to be merged. White scale bar equals 10 μm .

5.2.4.3 Immuno-Electron Microscopy for Localisation of DFK-12Myc

In an attempt to get an even finer resolution view of the location of the DFK-12Myc protein attempts were made to perform transmission electron microscopy on immunogold labelled sections of fixed parasites. As the 12Myc epitope tag was fused to the C-terminus of the DFK protein, directly adjacent to the protein kinase domain, it was hoped that this analysis would show which side of the cell membrane the protein kinase of DFK domain resides.

In the DFK-12Myc expressing cell line, STL2378, immunogold staining was detected on the cytosolic side of the flagellar membrane, the cell membrane and on the flagellum (Figure 5-9a-d). Staining was sometimes observed in the nuclei of the cells but, as this also occurred in 2T1 control samples, can be regarded as non-specific. Other non-specific staining was occasionally seen away from cells (Figure 5-9e,-f). Immunolabelling was weak in comparison to IFA analysis, but was consistent with previous findings and specific between test and control immuno-EM samples. Parasites were fixed with 2% formaldehyde/0.2% glutaraldehyde PBS to increase epitope preservation though this fixative appeared to be quite poor in many of the cells, leaving some artefacts (irregular, often round, less electron dense areas in the cytoplasm of cells). Despite this, the analysis is consistent with DFK localising to the cell membrane, with the protein kinase domain most likely to be intracellular.

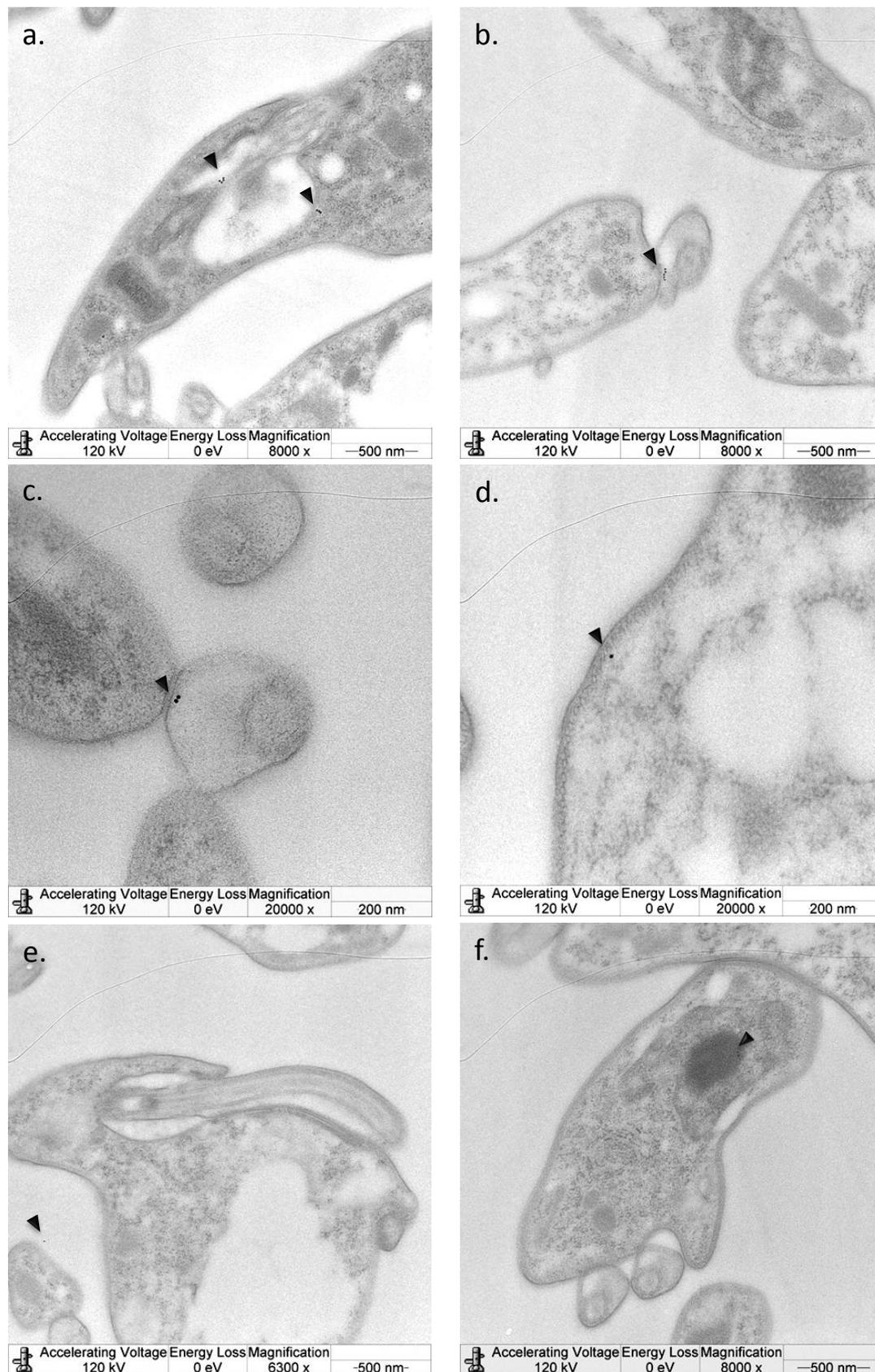


Figure 5-9: Transmission immune-electron micrographs from cell line STL2378 treated with mouse anti-Myc IgG and anti-mouse immunogold (10 nm) spheres. Immunogold labelling is highlighted by arrowheads. Panel a. shows immunogold labelling of the flagellum (inside the flagellar pocket) and internal membrane of the flagellar pocket. Panels b. and c. show intraflagellar membrane labelling. Panel d. shows labelling on the internal face of the cell membrane under the microtubule corset. Panel e. and f. depict some of the non-specific labelling seen in 2T1 cell control samples.

5.2.4.4 Immunoprecipitation of DFK-12Myc

As the anti-Myc antibody was clearly effective at detecting the epitope-tagged DFK protein attempts were made to immunoprecipitate DFK-12Myc from cell lysates. This was done with the aim of detecting interacting partners of DFK and so may provide information on how DFK effects its repression of the differentiation process. Immunoprecipitation (IP) was attempted with DynaBead technology (Life Technologies) and the Pierce c-Myc IP kit. The results of an IP conducted using the Pierce Kit (and its standard protocol) is presented in Figure 5-10. A large number (5×10^8) of cells were used to increase the chance of detecting weakly interacting proteins. A control immunoprecipitation was also performed using *T. brucei* 2T1 cells that do not express a Myc tagged protein. Elution fractions from DFK-12Myc samples and 2T1 samples contained proteins visible using Coomassie stained SDS-PAGE analysis (Figure 5-10a). Samples of total cell lysates, unbound, wash and elution fractions were taken for western blot to detect DFK-12Myc (Figure 5-10b). This showed that DFK-12Myc was expressed in the STL2378 cell line used in this assay and that it was capable of being immunoprecipitated, though a substantial amount remained in the unbound fraction. The western blot contains bands corresponding to the heavy and light chains of the co-eluted mouse anti-Myc IgG and also bands corresponding to breakdown products of DFK, despite the inclusion of protease inhibitors in the lysis buffers.

The IP elutions from tagged cell lines and 2T1 controls suggested a lot of non-specific binding was occurring with the agarose resin of the kit (Figure 5-10a). However, despite either unequal loading or elution in this example, it appeared that a band of DFK-12Myc was visible on a coomassie stained SDS-PAGE gel. Further to this initial test, a repeat of this experiment was performed with the eluates being sent to The University of Dundee for mass spectrometry analysis to investigate DFK for phosphorylation sites. This analysis did not detect any phosphopeptides corresponding to DFK (M Urbaniak, Unpublished).

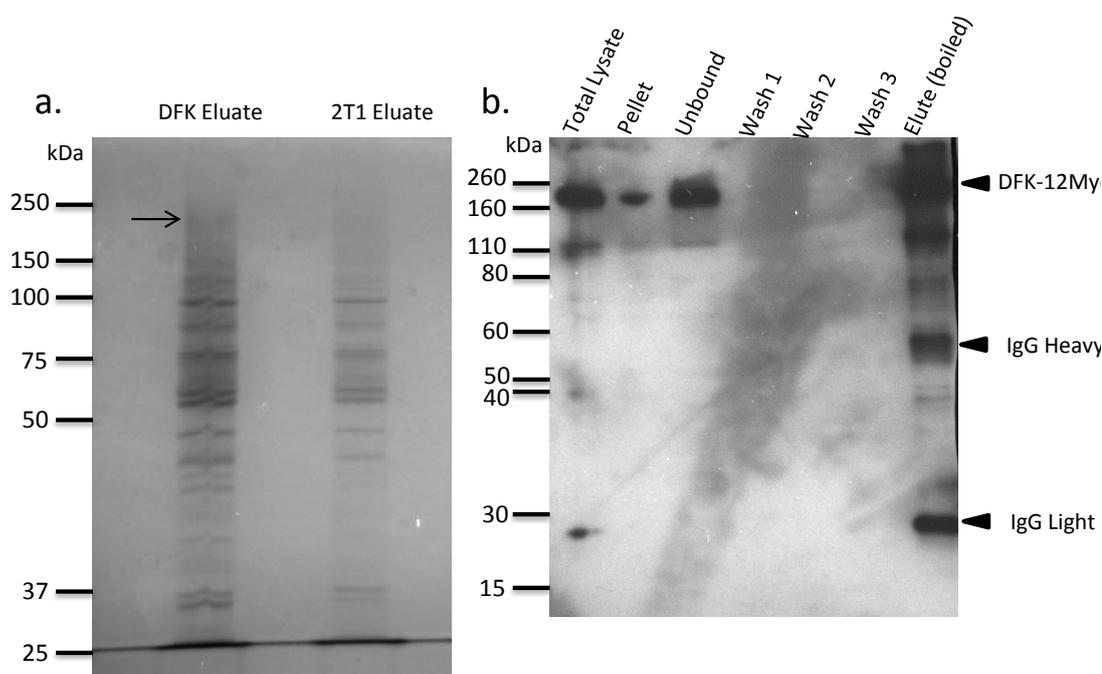


Figure 5-10: Immunoprecipitation of DFK-12Myc. Cell lysates from STL2378 (containing DFK-12Myc) and a *T. brucei* 2T1 negative control were incubated with anti-Myc antibody conjugated agarose beads (Pierce) and eluted by boiling in SDS-PAGE buffer. In panel a. the comparison to a control pulldown of *T. b. brucei* 2T1 lysate shows high levels of non-specific binding. A band potentially corresponding to DFK-12Myc is visible between 150 kDa and 250 kDa (Arrow), however loading is visibly not equal. Panel b. confirms immunoprecipitation of DFK-12myc by western blot using mouse anti-Myc IgG and anti-mouse HRP IgG, heavy and light chains from anti-Myc IgG liberated from agarose beads are visible (Arrow heads).

5.2.4.5 Transcriptomic analysis of DFK RNAi cell lines

The natural differentiation from BSF to PCF is associated with dramatic remodelling of the transcriptome and proteome (Jensen et al., 2009; Kabani et al., 2009; Queiroz et al., 2009; Kolev et al., 2010; Urbaniak et al., 2012a). We had determined that DFK RNAi was inducing EP procyclin expression, morphological changes and (unquantified) kinetoplast repositioning but had yet to detect other biological changes associated with differentiation. We therefore sought to analyse the process in an untargeted manner. The aim was to determine if the differentiation was representative of the natural differentiation process or was only the restructuring of the cell and its surface, without the commensurate changes in gene expression and metabolism etc. In order to do this we initially chose a transcriptomic analysis to identify global changes in RNA levels, in order to try and identify the gene expression signature of BSF-PCF differentiation.

Total RNA from DFK RNAi cell lines (STL201/STL202) was extracted from RNAi induced (48h) and uninduced cultures (2 bioreplicates), purified and prepared for RNA-seq at Glasgow Polyomics Facility. There it was prepared, and mRNA amplified using oligo-dT and multiplexed. This was then sequenced using an Illumina Genome Analyzer Iix. This generated paired-end reads (4 million per sample, with an average size of 110 bases), with an average insert size of 200 bases. GC content was 49% for all libraries. The read quality was controlled by trimming on mean quality score >Phred 20 (equating to <1% error rate). They were also checked for adapter or primer contamination. Non-redundant read data was aligned to the *T. b. brucei* 927 reference genome from TriTrypDB version 4.0 using Bowtie (0.12.7), and analysed for differential gene expression using Cufflinks (1.3.0). Cufflinks determined differentially regulated genes by calculating an uncorrected p-value for the fold change in reads for a gene and a q-value (a corrected p-value using a false discovery rate of 0.05). If the p-value is less than the q-value the gene is considered significantly differentially regulated. This analysis was performed by Dr Nick Dickens at the Wellcome Trust Centre for Molecular Parasitology.

The analysis of DFK RNAi lines identified a total of 479 significantly differentially regulated genes, of these 191 were identified individually in either of the bioreplicates. The remaining 288 were identified in both bioreplicates. Of the 479 genes identified as differentially regulated 242 were downregulated and 237 upregulated. These were compared to genes previously identified as differentially regulated in differentiating trypanosomes or PCF trypanosomes by microarray or RNA-seq (obtained from Jensen *et al.* 2010 and Siegal *et al.* 2010) using the search strategy functions of TriTrypDB. The genes used for differentiation time points came from those identified in Kabani *et al.* 2009 and Quiroz *et al.* 2009. Using the default settings of TriTrypDB the results of genes significantly up- or downregulated were pooled and then compared to the results of the DFK RNAi transcriptome. For early time points (representative of stumpy/early PCF) the reference samples were set to BSF form and compared to high density BSF, 0, 1 h and 1.5 h post CCA induced differentiation. For later time points this was set to compare BSF to 6 h-72 h post CCA induction of differentiation. The data comparing BSF to established PCF came from studies by Siegal *et al.* (2010) and Jensen *et al.* (2009) combined. The comparisons are

presented in Table 5-2, but overall when all the differentially regulated genes were tallied up from previous studies, 415 of the 479 genes detected in DFK RNAi induced cells were present in these studies. The full list is presented as an Excel Spreadsheet in supplemental file S3 on the appended CD-ROM. In order to determine if this was more than can be expected by chance a Chi-Squared test was performed against a list of 500 genes that should represent a random distribution over the genome. This showed that the genes found to be differentially regulated in DFK RNAi cell lines are significantly more likely to be found in other transcriptomes of differentiating trypanosomes than a random sample of genes. This is supportive of a true BSF-PCF differentiation process occurring after ablation of DFK.

Genes for certain processes were identified and exemplified (Table 5-3), based on previous findings in other transcriptomic analyses of differentiation (Jensen et al., 2009; Kabani et al., 2009; Queiroz et al., 2009). As the expression of EP procyclin had already been detected by IFA it was reassuring to find the transcripts for EP1, 2, 3 were all upregulated as well as for GPEET2. These transcripts were hugely upregulated; for EP2 the levels rose from 2.6 reads per sample to over 67 reads per sample - a fold change of 25.38. For GPEET2 a 45-fold increase was detected. If this is reflected at protein level assaying by FACS to detect GPEET expression may show an even higher proportion of differentiated cells than using EP procyclin as a marker for differentiation. Transcripts for VSG related genes were downregulated, whereas transcripts for TbPAD1 and TbPAD2 were upregulated. Evidence for mitochondrial development is supported by the increase in transcripts for cytochrome oxidase IV, VIII and X and putative NADH-cytochrome b5 reductases. There were changes in the levels of transcript for the RNA binding proteins, some of which are known to either be differentially regulated in the procyclic form or essential to them in RNAi analyses (Wurst et al., 2009). Taken together, these data suggest the formation of mature PCF parasites following DFK RNAi.

The differentiation phenotype (determined by expression of EP procyclin) following DFK ablation by RNAi only manifests in 5% of parasites at 48 h and 17% of parasites at 72 h therefore the changes detected in the DFK RNAi transcriptomes probably only represent the transcripts exhibiting the largest amount of change. Changes in low abundance transcripts may be masked by

transcript from the other 95-83% of cells not undergoing differentiation, assuming their transcriptome is not undergoing changes to which the cells are not responding. The fold change of the differentially regulated transcripts in the 5-17% of differentiating cells is therefore likely to be much higher than that detected by RNA-seq analysis. In this experiment, differentiation was quantified by expression of EP procyclin on the cell surface. It may be that the changes in mRNA levels are detectable to a greater extent than protein levels due to the time taken to synthesise EP procyclin and traffic it to the cell surface.

Table 5-2: Overview of genes detected as differentially regulated (up or down) in DFK RNAi cells and the existing data available in TriTrypDB. Following this is the output of the Chi-Squared test performed using Prism Graphpad.

Life cycle stage	Number of Genes					
	Differentiation early		Differentiation late		BSF vs PCF	
	UP	DOWN	UP	DOWN	UP	DOWN
TriTrypDB	610	517	650	551	2199	3156
Shared with DFK RNAi (479)	41	62	77	102	180	226

Chi-square

Chi-square, df

281.9, 1

P value

< 0.0001

P value summary

One- or two-sided

Two-sided

Statistically significant? (alpha<0.05)

Yes

Strength of association

Difference between proportions

Fraction of top, bottom row in left column

0.8664, 0.3400

Difference between fractions

0.5264

95% confidence interval of difference

0.4649 to 0.5879

Data analyzed

DFK

Diff Reg

415

Not Diff Reg

64

Total

479

Random

170

330

500

Total

585

394

979

Table 5-3: Selected transcripts identified as differentially regulated in DFK RNAi induced cells (48h), supporting the differentiation to PCF phenotype. Those marked with a * are fold changes for samples where no transcript was detected in the uninduced culture, to perform the fold change calculations read numbers were set to 1. The column labelled Transcriptome describes whether the transcript was detected in both bioreplicates (Common) or in only one of the cell lines (STL 201/202).

Category	Transcriptome	Gene ID	Gene Product	Fold Change
RNA Binding	Common	Tb11.01.3940	RNA-binding protein, putative (RBP9)	-4.70
	Common	Tb11.01.2580	pumillio/PUF RNA binding protein 11 (PUF11)	-2.64
	Common	Tb927.6.3480	RNA-binding protein, putative (DRBD5)	-1.98
Surface Modification	Common	Tb11.01.3915	RNA-binding protein, putative (RBP5)	4.99
	Common	Tb927.7.5930	Protein Associated with Differentiation (TbPAD1)	2.22
	Common	Tb927.7.5940	Protein Associated with Differentiation (TbPAD2)	3.70
	Common	Tb927.6.520	EP3-2 procyclin,PARP A-beta,surface protein EP3-3 procyclin precursor,procyclic form specific polypeptide A-beta precursor	20.54
	Common	Tb927.10.10260	EP1 procyclin (EP1)	20.21
	Common	Tb927.10.10250	EP2 procyclin (EP2)	25.38
	Common	Tb927.6.510	GPEET2 procyclin precursor,PARP A-alpha,procyclin A-alpha,procyclic form specific polypeptide A-alpha precursor	45.05
	Common	Tb927.6.480	EP3-2 procyclin,PARP A-beta,surface protein EP3-2,surface protein EP3-2 procyclin precursor,procyclic form specific polypeptide A-beta precursor	20*
	Common	Tb927.6.450	EP3-2 procyclin,PARP,procyclin PARP A,procyclin B1- alpha,procyclic acidic repetitive protein A.beta,procyclic form specific polypeptide B1-alpha precursor	15*
	Common	Tb927.10.11220	procyclic form surface glycoprotein (PSSA-2)	2.80

Glycolysis	Common	Tb927.8.7300	variant surface glycoprotein (VSG)-related, putative	-3.48
	Common	Tb927.8.7320	variant surface glycoprotein (VSG)-related, putative	-3.58
	Common	Tb927.3.1470	variant surface glycoprotein (VSG)-related, putative	-3.15
	Common	Tb927.3.1520	variant surface glycoprotein (VSG)-related, putative	-3.15
	Common	Tb927.1.5170	variant surface glycoprotein (VSG)-related, putative	-3.97
	Common	Tb11.02.1566	variant surface glycoprotein (VSG)-related, putative	-3.41
	Common	Tb927.5.130	variant surface glycoprotein (VSG)-related, putative	-3.42
	Common	Tb927.3.1510	variant surface glycoprotein (VSG)-related, putative	-2.98
	Common	Tb927.3.1500	variant surface glycoprotein (VSG)-related, putative	-3.32
	201-only	Tb927.1.700	phosphoglycerate kinase (PGKC)	-2.57
	Common	Tb927.1.3830	glucose-6-phosphate isomerase, glycosomal (PGI)	-2.61
	Common	Tb927.3.3270	ATP-dependent phosphofructokinase (TbPFK)	-3.53
	Common	Tb927.10.5620	fructose-bisphosphate aldolase, glycosomal (ALD)	-2.93
	Common	Tb927.10.14140	pyruvate kinase 1 (PYK1)	-3.58
Respiration	Common	Tb927.1.4100	cytochrome oxidase subunit IV (COXIV)	2.08
	Common	Tb927.4.4620	cytochrome oxidase subunit VIII (COXVIII)	4.13
	Common	Tb927.8.1890	cytochrome c1, heme protein, mitochondrial precursor	2.09
	Common	Tb11.02.1230	NADH-cytochrome b5 reductase, putative (B5R)	2.84
	Common	Tb11.01.7190	NADH-cytochrome b5 reductase, putative (B5R)	2.15
	Common	Tb11.01.4702	cytochrome oxidase subunit X (COXX)	3.40
	202-only	Tb11.01.3780	cytochrome oxidase assembly protein, putative	2.94
	202-only	Tb927.3.680	cytochrome P450, putative	2.00
	202-only	Tb927.7.2700	NADH-cytochrome b5 reductase, putative (B5R)	1.91
	201-only	Tb927.10.7090	alternative oxidase (AOX)	-2.31

5.3 Discussion

The roles of protein kinases and phosphatases in the various stages of trypanosome development and differentiation are becoming more apparent and increasingly more defined. Several protein kinases (ZFK, MAPK5) have been shown to repress the differentiation of slender to stumpy forms at parasitaemia densities which would be too low (Vassella et al., 2001; Domenicali Pfister et al., 2006b). TOR4 has been shown to retain slender BSF in that state and depletion of TOR4 leads to irreversible commitment to stumpy forms (Barquilla et al., 2012). Once in this form, TbPTP1 and TbPIP39 combine to hold the parasite in G₀ arrest until the addition of CCA inactivates TbPTP1 allowing TbPIP39 to trigger stumpy to PCF differentiation following its entry into glycosomes in its phosphorylated, active form (Szöör et al., 2006, 2010). Following this differentiation event TbMAPK2 is essential in driving the proliferation of PCF parasites (Müller et al., 2002a). In this study we implicate another protein kinase as a negative regulator of the differentiation process. RNAi knockdown of DFK led to spontaneous differentiation of 17% of a monomorphic BSF population to PCF parasites. This phenotype is reminiscent of the RNAi of TbPTP1 in monomorphic cells or the inhibition of TbPTP1 using BZ3 suggesting DFK may act somewhere in the same pathway (Szöör et al., 2006). In order to probe the function of DFK a number of approaches were taken to determine the localisation of the protein and the process of DFK RNAi induced differentiation. These approaches suggested DFK is an integral cell membrane protein, and that once it is removed from BSF cells by RNAi the differentiation process reflects the natural BSF-PCF differentiation process. Thus it appears the DFK acts as a negative regulator of BSF-PCF differentiation.

5.3.1 Confirming the DFK Phenotype

The DFK RNAi phenotype was detected by visual observation of RNAi induced cultures during the pTL screen of the trypanosome kinome for potential drug targets. Having observed a visible morphological phenotype apparently consistent with PCF parasites the cell line was assessed further in order to determine if these cells had other markers that have been used to define the PCF stage. Several results are suggestive of this and some appear to strongly confirm it.

When RNAi of DFK was induced a growth defect in cultured parasites was detected by haemocytometry. The growth defect was more severe than the one exhibited by NEK 16 (Figure 3-5) and should have been detectable by the alamar blue assay. After 72 h induction of DFK RNAi there were fewer cells than in the comparable NEK 16 sample (which did generate a detectable alamar blue phenotype). The smaller number of DFK RNAi induced cells were therefore reducing resazurin to resorufin at the same rate as the DFK RNAi uninduced samples. Thus they generated the same fluorescence levels, giving an alamar blue assay ratio of >0.9 and so not placing DFK in the loss of fitness group. One explanation for this could be that despite there being reduced cell numbers in the alamar blue assay, these cells were more metabolically active (as may be expected if the cells were developing their mitochondrion, as occurs in stumpy or PCF parasites). Conversely, the NEK 16 RNAi may lead to a decrease in the ability to reduce resazurin but only generate a small growth defect.

It was observed that 10% of the cells exhibited the morphology of PCF parasites and were tested for other markers of this life cycle stage. The cells were tested for the presence of EP procyclin - which is usually a marker for late procyclic form cells (Acosta-Serrano et al., 2001). This was visible on individual cells using IFA analysis and was an amenable phenotype for assessment using a FACS assay. Analysis of EP procyclin expression in RNAi induced cultures showed that 17% were EP procyclin positive by 72 h. This was confirmed by initial experiments using an IN-Cell 2000, which showed 20% of cells were EP procyclin positive.

Although cells expressing EP procyclin were also observed to reposition their kinetoplast this remains to be thoroughly quantified and further experiments are necessary to confirm the morphological changes in the RNAi induced parasites.

5.3.1.1 Testing for Mitochondrial Development Using Respiratory Inhibitors

In an effort to determine whether the PCF-like cells were developing a fully functional mitochondrion as would be expected for PCF trypanosomes, an approach was attempted to use respiratory inhibitors to enrich or reduce for PCF cells in the population. BSF trypanosomes possess only one terminal alternative oxidase (TAO) in the mitochondrion which does not contain cytochrome oxidase (COX). Upon differentiation to the PCF form the cytochrome oxidase system is

expressed and respiration mostly occurs through the COX system, though TAO is still functional - TAO can be selectively inhibited with SHAM (salicylhydroxamic acid) and COX with KCN (potassium cyanide)(reviewed in (Chaudhuri et al., 2006)). In order to infer whether this was happening in the DFK RNAi induced cell lines they were treated with KCN and SHAM. KCN treatment was expected to block the development of the PCF-like cells if they were developing a complex mitochondrion containing COX - which would be stronger evidence of a true BSF to PCF differentiation. Using the previously published protocol of Le Febvre and Hill, DFK RNAi induced and uninduced cultures (and *T. brucei* 927 PCF controls) were exposed to varying concentrations of respiratory inhibitors (KCN: 100 μ M, SHAM: 250 μ M)(Le Febvre and Hill, 1986). It appeared that the KCN treatment had only a very slight effect on established PCF parasites or the DFK RNAi induced cultures at any concentration. SHAM had a toxic effect on all samples, it proved slightly more toxic to DFK RNAi induced cells due to their inherent growth defect. At the end of the initial test experiment too few cells remained to perform follow up experiments. The 2T1 cell line did not respond to the compounds in the same manner as previous studies(Le Febvre and Hill, 1986). Perhaps a more simple approach should have been undertaken to look for visible increases in the size of the mitochondrion after labelling with a dye such as MitoTracker or TMRM - which would have allowed qualitative analysis by microscopy and quantification of any change in fluorescence in a FACS assay, along with the expression of EP procyclin, to analyse the population of DFK RNAi induced cultures. The NADH diaphorase assay could also have been used to determine the phenotype of mitochondrial development (Vickerman, 1965). These experiments were superseded by the transcriptomic data showing increase in expression levels of several cytochrome oxidase subunits, though it would be interesting to know if they are expressed in a functioning mitochondrion.

5.3.1.2 Transcriptomic Analysis

Although the observable changes in the cells suggested a PCF phenotype the opportunity to perform a transcriptomic analysis of the population presented a chance to explore this in more detail. The transcriptomic data appear to show that the populations of cells undergoing DFK RNAi generated an expression profile that is very suggestive of a differentiation to a PCF cell. Though the number of genes detected as being differentially regulated in DFK RNAi lines is

much smaller than those detected in other studies comparing efficient, synchronous differentiation there are signatures that are distinctive of stumpy and PCF trypanosomes. This may be explained by the fact that only 5% of the cells appear to differentiate at 48 h, the time of RNA extraction, so there may be a population of stumpy* cells and a population of cells undergoing stumpy to PCF differentiation. As there was a limit to the amount of samples that could be analysed (due to cost, time and ability to multiplex the assay) a timepoint was chosen that was expected to observe cells going through the differentiation process, rather than being at the end of it. Whether this readjustment of the transcriptome is represented in all the cells in the population is undetermined. One would expect that as only a small proportion of cells in the population differentiate after DFK RNAi that the transcriptome remodelling is only occurring in these cells, and this may lead to masking of changes in these cells by the majority of unchanged cells. It would be difficult to expect large changes in the transcriptome of a cell to not lead to phenotypic changes in the cell. Because the process of differentiation post-DFK ablation is partial and also asynchronous this presumably dampens the signal coming from differential gene expression (in the differentiating cells), and could explain the low number of genes detected compared to other studies that examined either synchronous or stable populations. However, the detection of key markers (Table 5-3) is enough to give a good idea that the differentiation phenotype is real, especially when combined with other evidence such as detection of EP procyclin protein expression.

Tellingly, the transcripts for PAD1 and PAD2 were detected as upregulated; a result which has since been confirmed for the case of PAD2 by qPCR (M. Saldivia, Unpublished). Though PAD2 can be upregulated very rapidly by cold shock, PAD1 is a specific marker for stumpy form trypanosomes and the expression of this could suggest the DFK RNAi lines are progressing through a stumpy or stumpy* phase during the differentiation to PCF. However, whether this is expressed or trafficked to the surface of the cell requires further investigation by western blot or IFA for confirmation. PAD1 is upregulated at mRNA level in PCF versus BSF cells even though the protein is only detectable in stumpy forms (Dean et al., 2009). VSG and VSG-related transcripts are downregulated in DFK RNAi lines, and this has also been confirmed by qPCR (for VSG 221) (M Saldivia, Unpublished). This suggests that RNAi of DFK is triggering the formation of

stumpy-like cells which proceed to PCF cells in the absence of a stimulus such as cis-aconitate.

During the natural slender-stumpy-PCF differentiation process the DFK transcript is detected in two expression profiling studies to be downregulated (Figure 5-11, TriTrypDB). Two other studies did not show the same magnitude of downregulation though there was a slight downward trend. The study by Kabani *et al.* 2009 used 5 bioreplicates, strengthening the reliability of the data from this study. This suggests that DFK is important for keeping BSF forms in that lifecycle stage but is less important or not required for the PCF stage, and its expression is reduced once it has served its purpose in the BSF. DFK was not detected in several recent proteomic experiments that compare BSF to PCF stages (Gunasekera *et al.*, 2012; Urbaniak *et al.*, 2012a), which if it had would possibly offer more robust evidence for this idea.

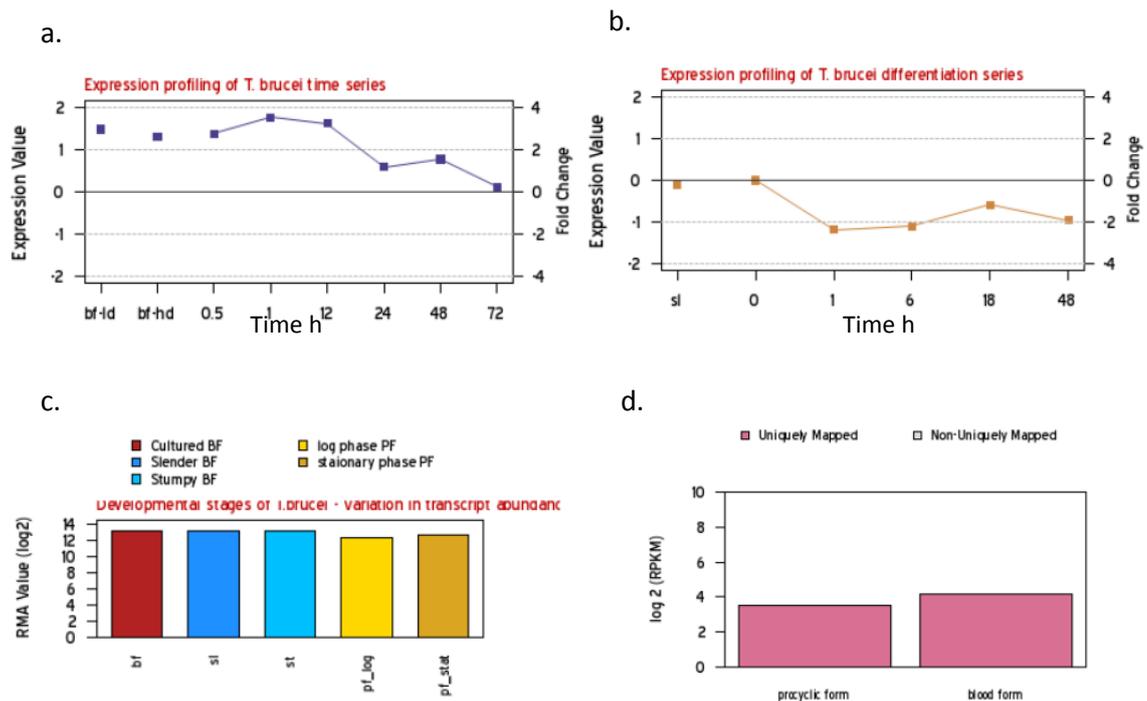


Figure 5-11: Transcript expression profiling of DFK from a. Quieroz *et al.* 2009, b. Kabani *et al.* 2009 c. Jensen *et al.* 2009 and Siegel *et al.* 2010. The graphs in a. and b. show the level of DFK transcript in slender BSF parasites and over the synchronous differentiation of stumpy form parasites to PCF parasites. Panel a. used cells derived from axenic cultures whilst b. and c. used stumpy forms derived from animal infections. Study d. used established cultures of BSF and PCF forms. Adapted from TriTrypDB.

In overview, the transcriptome of DFK RNAi mutants shows clear signs of differentiation - but it probably does not provide information on the mechanisms

causing that differentiation. Changes in the transcriptome in this instance probably occur as a response to changes in other pathways/process after ablation of DFK. It is difficult to see how this analysis assists in determining the pathways that DFK is mediating, though it does give a clearer picture of the nature of the phenotype. RNAi of DFK does not just lead to induction of procyclin expression but of a wide range of processes and consequences for the cells, culminating in an inefficient and fatal differentiation. Ultimately, whether this is due to ablation of DFK activity or simply by removal of a protein or disruption of a protein complex remains to be seen in future experiments.

5.3.2 Localisation and Topology of DFK

The location of a protein kinase can give clues to the function it performs. The case for DFK being located in the cell membrane appears quite strong. Prior to this work, DFK had been detected in a study examining the proteome of the flagellum - which had sought to resolve flagellar membrane proteins from flagellar matrix proteins. Oberholzer *et al.* detected DFK in the membrane fraction of the flagellum, and as it is predicted to have multiple TM domains, this would appear to be a sensible localisation (Oberholzer *et al.*, 2011). Their study did not confirm the localisation of DFK by another method, and even goes as far to suggest that DFK is a receptor kinase. This is an interesting concept but no exploration of this suggestion has yet been published. In the present study, we determined that DFK localised to a membranous fraction of cell lysates. This was a simple fractionation so it could not be determined whether the protein was in a cell membrane or an organellar membrane fraction. However, when the epitope tagged version of DFK was localised by microscopy it appeared to reside over the entire cell membrane, including the flagellar membrane.

Different TM domain prediction algorithms produced varying results for the number of TM domains in DFK. One would expect a “receptor kinase” to possess an intracellular kinase domain to perform phosphorylation of the next protein in a signalling cascade or of an effector protein. The 12Myc epitope tag is located at the C-terminus of the protein, effectively fused to the kinase domain. The IFA staining pattern on cells with a non-permeabilised cell membrane would suggest the epitope tag, and thus the kinase domain, is extracellular - if correct this may be considered an unconventional arrangement. However, when these samples

were prepared for microscopy they were first dried onto poly-L-Lysine treated multi-well slides to adhere them. This probably compromised the cell membrane, thereby allowing antibodies to permeate through the cell membrane. It is therefore difficult to assess from these data whether the protein kinase domain is intra- or extracellular. Certainly when the membrane is permeabilised using Triton X-100 the signal from anti-Myc IFA decreases and becomes more punctuate, suggesting that DFK-12Myc is being stripped away with the membrane. Further analysis by immunoelectron microscopy does point to a cell membrane localisation of DFK, and immunogold labelling consistently appears on the internal face of the cell membrane in the DFK-12Myc cell line, suggesting that the epitope tagged protein kinase domain of DFK-12Myc is intracellular. Though the amount of staining is weak, possibly due to the age of the anti-mouse immunogold reagent, it is consistent - and is consistent with other data on the localisation of DFK. Repeats of this experiment with optimised reagents may provide a clearer picture of the localisation of DFK.

Further IFA analysis could assess the localisation of the DFK kinase domain further by treating cells with a protease to remove proteins on the cell surface. If the DFK-12Myc signal is ablated it would confirm an extracellular localisation of the kinase domain, if not it would confirm an intracellular localisation. Further optimisation of the immune-EM procedure would probably be the ideal way of determining the localisation of the protein kinase domain, and this combined with deletion mutants with epitope tags following the various predicted TM domains could potentially be used to determine the topology of DFK in the cell membrane. The current problems with the immune-EM analysis are probably due to the age of the anti-mouse immunogold, which appeared to be heavily precipitated, leading to a low concentration of the reagent in the supernatant - unfortunately this could not be redressed due to time constraints and limited access to the Integrated Microscopy Facility.

5.3.3 Functional Role of DFK

Currently, in *T. brucei*, there are no defined protein kinase signalling pathways - although a protein phosphatase cascade has been described - coincidentally acting in the BSF-PCF differentiation process. DFK appears to be a candidate ripe for investigation due to the striking phenotype exhibited after its depletion.

DFK is predicted to contain a protein kinase domain most similar to a STE11 family kinase, also known as a MAPKKK. Using the example of the yeast mating pheromone response, these typically sit at the top of a MAPK signalling cascade; after being activated by STE20 (MAPKKK). MAPKKs, such as STE11, will typically phosphorylate a MAPKK, such as STE7, which will then go on to activate a MAPK such as Kss1 or Fus3. MAPK signalling components are usually organised together at the plasma membrane by means of a scaffold protein, such as STE5 or Pbs2 in yeast or KSR, IQGAP or β -arrestin in mammalian cells (Qi and Elion, 2005; Brown and Sacks, 2009; Good et al., 2011). As signalling cascades in *T. brucei* are completely undefined, no scaffold proteins supporting signalling components are known either. The future experiments to perform SILAC-labelled quantitative MS/MS on DFK-12Myc IP samples may provide information on whether other potential MAPK pathway proteins associate with DFK and whether this type of signalling complex exists. No other protein kinase demonstrated this phenotype during RNAi, though this does not rule out other MAPKKs and MAPKs being involved in this process until the efficacy of knockdown has been performed for all cell lines. The possibility of functional redundancy further along in the pathway (which is evident in well characterised yeast MAPK pathways (Wang et al., 2011)) exists and this may explain why the DFK RNAi phenotype was not observed for any other protein kinase RNAi.

If DFK is a MAPKKK it would be atypical as it is an integral membrane protein. The architecture of the protein domains is very reminiscent of a transmembrane receptor kinase. It possesses multiple transmembrane domains, although the topology of the protein is still unclear despite the current analysis. Since the determination of structures of a large number of membrane proteins it is clear that they possess more topological diversity than first assumed, i.e. the length and angle of the transmembrane helix across the membrane can vary significantly. Predicting topology by bioinformatic methods should also be done with caution (Von Heijne, 2006). One of the predicted transmembrane regions is detected by PFAM analysis as a HAMP domain. These were originally defined in prokaryotes and are common to a multitude of prokaryotic transmembrane receptor histidine kinase and chemotaxis receptors (reviewed in (Parkinson, 2010)). They are usually confined to regions near the cytoplasmic side of the membrane and consist of a sequence approximately 50 amino acids in length

forming two amphipathic helices connected by a linker region. When extracellular receptor domains are bound by their ligand, conformational changes occurring in this region can be transmitted via a TM domain to the HAMP domain which can either tighten or loosen. This change can induce kinase-on/off states in the relevant domain of the protein. Many of the chemoreceptors and sensor kinases possessing these domains operate as homodimers in the cell membrane. Another interesting feature detected by PFAM search are the TPR2 repeats which represent a motif that allows protein-protein interaction. It could be speculated that this assists other proteins to complex with DFK or even allows dimerisation of DFK to occur. This conjecture is totally untested but aspects of the DFK protein architecture do at least hint of these possibilities.

Searching for DFK in OrthoMCL shows that it is confined to kinetoplastids. TbDFK is predicted to have both a syntenic and non-syntenic orthologue in *Leishmania braziliensis*, *L. infantum*, *L. major*, *L. mexicana* and *L. tarentolae*; however, only the syntenic orthologues provide the best sequence orthologues. The two orthologues in *T. cruzi* (as predicted by OrthoMCL) appear to align poorly to the *T. brucei* DFK sequence and the closest match at the level of protein sequence is the gene TcCLB.511727.210. There is a syntenic orthologue in *T. b. gambiense*. *Trypanosoma congolense* does not contain an orthologue of DFK and *T. vivax* is predicted to contain two gene fragments with orthology to DFK - though this may be a sequencing or assembly error in this genome. Performing a tblastn search of the *T. evansi* and *T. equiperdum* genome detect a gene identical to the *T. brucei* 927 DFK gene in each (D. Barry, Unpublished). Thus DFK appears to be present in organisms that do not, or in the case of *T. evansi*, cannot form procyclic forms (Brun et al., 1998).

Table 5-4: Sequence orthologues of TbDFK from published genomes. * denotes syntenic orthologue.

Species	Gene ID
<i>T. b. brucei</i> 927	Tb11.01.5650*
<i>T. b. brucei</i> 427	Tb427tmp.1.5650*
<i>T.b. gambiense</i>	Tbg972.11.15700*
<i>T. vivax</i>	TvY486_0044340,TvY486_0044360
<i>T. cruzi</i>	Tc00.1047053511727.210
<i>L. major</i>	LmjF.32.0810*
<i>L. infantum</i>	LinJ.32.0860*
<i>L. braziliensis</i>	LbrM.32.0900*
<i>L. mexicana</i>	LmxM.31.0810*
<i>L. tarentolae</i>	LtaP32.0880*

5.3.3.1 Testing for Effects of pCPT-cAMP and BZ3 on DFK RNAi lines

In order to try and determine at which stage in the differentiation process DFK operates at, attempts were made to modulate other parts of the pathway using small molecules. An attempt to enrich an uninduced population of the DFK RNAi line for stumpy* parasites using a cyclic-AMP analogue was made, as it was expected that if DFK was operating to induce stumpy-PCF differentiation then an increase in EP procyclin positive cells would be observed if stumpy* parasites were increased. If DFK was involved in mediating slender to stumpy differentiation and the PCF differentiation was an artefact of this then enriching for stumpy* parasites with cAMP-analogues might not have an effect on the proportion differentiating to PCF. Due to the ability of cAMP breakdown products to increase the proportion of stumpy form cells in the population it was attempted to test the effect 8-(4-chlorophenylthio)-cAMP (pCPT-cAMP), a cell permeable cAMP analogue on DFK RNAi cell lines (Vassella et al., 1997). This was expected to increase the proportion of stumpy* cells in the monomorphic 2T1 cell line. The aim of performing this was to then induce RNAi of DFK and observe any increase in the efficiency of differentiation following ablation of DFK, similar to RNAi of TbPTP1 after pCPT-cAMP treatment (Szöör et al., 2006). Trial treatments of DFK RNAi lines with 100, 500 and 1000 μ M appeared to be highly toxic to 2T1 derived cell lines- after an overnight incubation with this compound

the cells looked enlarged and unhealthy. RNAi of DFK was not attempted. This would be an interesting experiment to pursue under optimised conditions.

Attempts were also made to test what, if any, effect BZ3 had on the proportion of cells differentiating during DFK RNAi. This was done to determine if inhibition of TbPTP1 and removal of DFK generated additive or synergistic effects on the number of cells differentiating from BSF to PCF. Unfortunately both BZ3 also proved toxic to 2T1 cells at the concentrations used successfully in other studies (Szöör et al., 2006). Optimisation of these experiments could allow potentially informative data to be generated, otherwise RNAi vectors could be made to target multiple genes (e.g. DFK+TbPTP1) in an effort to circumvent the problems of inhibitor toxicity. Other proteins involved in aspects of differentiation, such as TOR4, do not yet have specific inhibitors developed or identified for use as research tools and double RNAi of these genes may be the only way to study their interactions.

5.3.4 Summary and Future Direction

DFK appears to be a repressor of differentiation in BSF cells to prevent them becoming PCF cells under inappropriate conditions. The pathway it participates in is as yet undetermined. DFK RNAi to 17% of cells differentiating to PCF as determined by expression of EP procyclin. Also, stumpy form specific transcripts were detected. However, detection of these at protein level would be required to determine the presence of stumpy cells categorically. It is likely that, due to the proportion of cells differentiating, DFK acts in procyclic cells as a molecular brake - similar to TbPTP1 (Figure 5-12). DFK protein has multiple transmembrane domains and features reminiscent of prokaryotic receptor kinases- further investigations should attempt to determine if DFK acts as a receptor, possibly entering a kinase-off conformation after encountering a stimulus. Interestingly, TbPIP39 triggers differentiation when entering the glycosomes in an (phosphorylated) active mode, whatever its substrate is must be phosphorylated by an as yet undetermined protein kinase: could the substrate of DFK also be trafficked to the glycosomes where it represses differentiation until dephosphorylated by TbPIP39? Removal of DFK by RNAi could simulate this situation and is one line of conjecture for the function of this protein kinase.

Clearly, much work remains to be undertaken before these answers will become apparent.

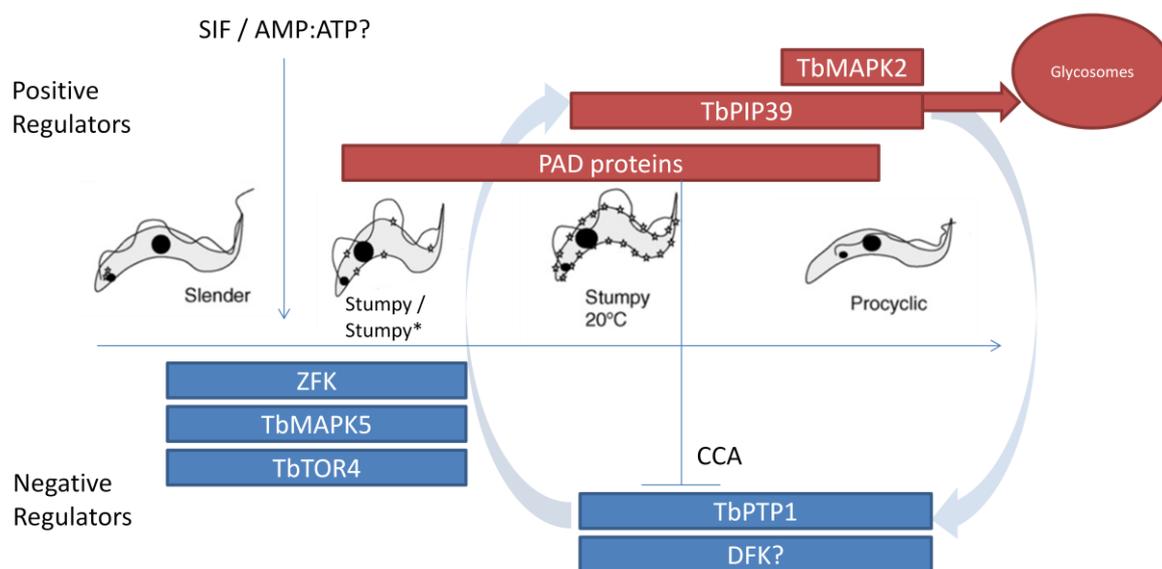


Figure 5-12: Schematic representing the known regulators of BSF to PCF differentiation. Known interactions are connected by arrows.

In order to advance the understanding of DFK's role in the differentiation process a number of lines of investigation can be taken. The search for a specific inhibitor of DFK is one such enquiry that is currently ongoing. As the spontaneous differentiation under RNAi was only observed for DFK it was reasoned that a specific inhibitor will also induce this phenotype (if DFK kinase activity alone is negatively regulating differentiation). A high-content screen has been developed to apply protein kinase focused inhibitor libraries to BSF trypanosomes and assess for EP procyclin expression after 72h (Figure 5-6c). Several compound libraries have been obtained from Roche, GSK and SULSA totalling over 6000 compounds, with the screen being conducted at the Scottish Bioscreening Facility - based at the University of Glasgow. Should a protein kinase inhibitor-like molecule be found to induce BSF-PCF differentiation it is probable that DFK would be the target, though further confirmation of this would be required to validate it fully - perhaps by *in vitro* assays and chemical genetic studies of DFK. Should an inhibitor of DFK be identified it would potentially allow for phosphoproteomic experiments to determine downstream phosphorylation events dependent on the presence of active DFK. Also, any inhibitors identified may provide a starting point for optimisation of drug-like compounds for use against DFK in a therapeutic manner. Trypanosomes expressing procyclin are

susceptible to complement mediated lysis by the alternative pathway and thus any differentiation of parasites by DFK inhibition would lead to their killing (Ferrante and Allison, 1983). This would explain why no EP-procyclicin expressing parasites were detected in the mouse infections after DFK RNAi induction.

Work to completely genetically validate DFK as playing a role in this pathway must be undertaken. Re-expression of recoded, RNAi refractory mutants has been attempted but the generation of the expression constructs using Gateway cloning has proved problematic. A restriction enzyme adapted vector, containing an active, recoded DFK gene, for performing these experiments is being synthesised. This will then be mutated at the catalytic lysine (K950M) to generate a kinase dead mutant. By using these two mutants in a DFK RNAi background it should be possible to determine the importance of DFK activity, and not simply presence of the protein, for repression of differentiation. Ideally this would have been undertaken before HCS screens for DFK inhibitors to confirm that DFK is an active protein kinase *in vivo*. If a potential DFK inhibitor is found during the HCS screen then this re-expression technique could be used to validate the inhibitor:target using chemical genetics to mutate residues important for inhibitor binding (and not for enzyme function) and observe for resistance to the inhibitor. Overexpression of a kinase dead, kinase domain might be another way of testing this, if a dominant negative phenotype can be achieved by the level of overexpression.

DFK presents an interesting opportunity to study trypanosomal signalling in the differentiation process due to its potential to interact with known modulators of trypanosome differentiation. It may be possible for this protein kinase to integrate environmental signals into signalling pathways allowing trypanosomes to respond to environmental changes and this would be a key hypothesis to investigate in the future. Also, it could be considered a novel drug target and as protein kinases are known to be druggable then the potential for specific inhibition of DFK must also be investigated further in order to further new opportunities for developing potential treatments of a widespread human and animal pathogen.

6 Concluding Remarks

RNA interference was discovered in trypanosomes in 1998 and since then the development and application of heritable, inducible RNAi systems has greatly facilitated the study of gene function in these parasites. As these technologies evolved they improved in terms of finesse (i.e. use of endogenous promoters instead of the T7 promoter), efficiency and reliability (i.e. using tagged loci). The development of the pTL system for rapidly generating stem-loop RNAi constructs has allowed the generation of several hundred RNAi plasmids, including a library to target the complement of the trypanosome kinome. The pTL system was based on an existing system for use in BSF parasites allowing efficient transfection into a single rRNA spacer locus, thus generating reliable RNAi performance. The speed and reliability of this system should facilitate the investigation of other gene families in *T. brucei*. In this study the system was used to interrogate the kinome of *T. brucei* for protein kinases that were important for the growth and replication of the clinically relevant bloodstream form parasite in order to identify potential drug targets. Several of the protein kinases identified as being essential to BSF parasites were further investigated as potential drug targets. In future studies the pTL and STL libraries could be applied to answer various biological questions. As protein kinases are implicated in regulating almost all cellular processes at some level the ability to test the effect of each *T. brucei* protein kinase individually on any given process makes these libraries a powerful resource.

Protein kinases have been investigated prodigiously as drug targets in the field of oncology and the application of this strategy to infectious disease is conceptually uncomplicated, as one intends only to stop a rapidly proliferating cell type. Candidate approaches have identified *T. brucei* protein kinases as drug targets but due to the high failure rates of drug development programmes the need to provide a large number of validated targets into pipelines means that genetic validation techniques must become more rapid. The pTL screen has partially genetically validated 53 protein kinases as drug targets and now defines where future validation efforts should be directed. This scope can be further refined by choosing protein kinases that might be amenable to work with, such as those that do not require co-factors such as cyclins or calmodulin for activity, or those that are likely to require activation by upstream kinases (e.g. MAPKs). This

would attempt to stack the odds in favour of progressing a target protein kinase some way through the drug development pipeline by increasing the likelihood of generating an *in vitro* activity assay successfully.

In particular this study targeted several families of protein kinase, the NEK and Orphan families, which displayed features predicted to make them attractive drug targets (if they were essential to the parasite). These families were among the first to be tested in the RNAi screen and a concurrent effort to express recombinant protein for a number of these was also performed. Those protein kinases which were essential to the parasite and also could be expressed recombinantly were prioritised for further assessment as drug targets. NEK 12.1 and NEK 12.2 were shown to be essential to the parasite under double-knockdown conditions. Further experiments are needed to assess whether these proteins are individually essential to the parasite but if NEK 12.1 activity can be fully genetically validated as individually essential then it could be further pursued as a drug target due to the ability to express it as an active recombinant protein that can be used in biochemical assays. With optimisation of the format, an HTS assay could be developed. The structural features NEK 12.1 is predicted to contain suggest it may be amenable for rational design of selective inhibitors. In theory this raises the possibility that these could be developed into therapeutic inhibitors. However, any selective inhibitor identified could potentially be used as a research tool to investigate the role of NEK 12.1 in trypanosome biology.

The characterisation of signalling pathways in *T. brucei* lags behind that of other model organisms. The signalling processes surrounding BSF to PCF differentiation represents those best characterised in the parasite. Identification of DFK as a negative regulator of this process adds a new facet of differentiation for investigation. Integration of the pTL library with assays to determine the phosphorylation status of other protein kinases and phosphatases may be able to determine the signalling events that precede differentiation. An effort to identify phosphosites on signalling or effector proteins and generate specific antibodies against these could then be used in systematic profiling of the library to identify those protein kinases that have an effect on that given residue. Clearly, this is a serious undertaking, mainly in the need to identify and understand the important phosphosites in members of a signalling cascade.

Therefore other processes maybe less complicated to investigate due to the ability to use reporter assays. Some biological processes may be more amenable to the application of reporter assays than others, but the ability to generate reporter cell lines is relatively straightforward. The majority of STL cell lines are compatible with any construct using a blasticidin or puromycin selectable markers allowing proteins to be tagged with fluorescent proteins or epitope tags. This flexibility will allow the generation of cell lines compatible with techniques to allow phenotypic screens to identify the effect of a kinase RNAi on a tagged protein or process. This has already been used in this laboratory for several proteins demonstrating the utility of such a library, and that it can be applied to more complicated questions than the identification of protein kinases required for cell viability.

RNAi of certain protein kinases can produce very defined phenotypes; for example, differentiation to PCF under DFK ablation, which was identified in the initial pTL screen. These specific or characteristic phenotypes present opportunities to develop assays to identify inhibitors that phenocopy the RNAi of these protein kinases. High-Content Screens are under development to identify inhibitors that phenocopy the DFK-RNAi-induced differentiation, and it is expected that these assays could be applied to other processes. Inhibitors that phenocopy a given RNAi are likely to target the same protein or pathway involved in that process and could be used in further studies investigating that process. These investigations could lead to new research tools, lead compounds for drug development, and will contribute to the understanding of trypanosome biology.

The protein kinase complement of *T. brucei* represents a richness of signalling interactions that have barely been investigated by the standards of some other eukaryotes. Cell signalling by reversible phosphorylation has been well studied in many other model organisms, where the ready availability of appropriate reagents provides the ability to investigate and understand pathways. A number of protein kinases have been investigated individually in *T. brucei* but these studies represent islands of data, with little interconnectivity. Given the tractability of *T. brucei* as a model organism the development of reagents and assays compatible with the pTL/STL libraries will make it possible to start to piece together the effect of protein kinases on signalling pathways and

other biological processes in this. This will contribute to our understanding of the basic biology of an early branching eukaryote that is also an important human and animal pathogen. By studying these processes we will generate opportunities for translational science to investigate new therapeutic avenues with the potential to positively impact on this neglected tropical disease.

List of References

Acosta-Serrano, A., Vassella, E., Liniger, M., Kunz Renggli, C., Brun, R., Roditi, I., and Englund, P.T. (2001). The surface coat of procyclic *Trypanosoma brucei*: programmed expression and proteolytic cleavage of procyclin in the tsetse fly. *Proceedings of the National Academy of Sciences of the United States of America* *98*, 1513–1518.

Adams, J.A. (2001). Kinetic and catalytic mechanisms of protein kinases. *Chemical Reviews* *101*, 2271–2290.

Alibu, V.P., Storm, L., Haile, S., Clayton, C., and Horn, D. (2005). A doubly inducible system for RNA interference and rapid RNAi plasmid construction in *Trypanosoma brucei*. *Molecular and Biochemical Parasitology* *139*, 75–82.

Allan, R.K., and Ratajczak, T. (2011). Versatile TPR domains accommodate different modes of target protein recognition and function. *Cell Stress & Chaperones* *16*, 353–367.

Alsford, S., Eckert, S., Baker, N., Glover, L., Sanchez-Flores, A., Leung, K.F., Turner, D.J., Field, M.C., Berriman, M., and Horn, D. (2012). High-throughput decoding of antitrypanosomal drug efficacy and resistance. *Nature* *482*, 232–236.

Alsford, S., and Horn, D. (2008). Single-locus targeting constructs for reliable regulated RNAi and transgene expression in *Trypanosoma brucei*. *Molecular and Biochemical Parasitology* *161*, 76–79.

Alsford, S., Kawahara, T., Glover, L., and Horn, D. (2005). Tagging a *T. brucei* RRNA locus improves stable transfection efficiency and circumvents inducible expression position effects. *Molecular and Biochemical Parasitology* *144*, 142–148.

Alsford, S., Turner, D.J., Obado, S.O., Sanchez-Flores, A., Glover, L., Berriman, M., Hertz-Fowler, C., and Horn, D. (2011). High-throughput phenotyping using parallel sequencing of RNA interference targets in the African trypanosome. *Genome Research* *21*, 915–924.

Amiguet-Vercher, A., Pérez-Morga, D., Pays, A., Poelvoorde, P., Van Xong, H., Tebabi, P., Vanhamme, L., and Pays, E. (2004). Loss of the mono-allelic control of the VSG expression sites during the development of *Trypanosoma brucei* in the bloodstream. *Molecular Microbiology* *51*, 1577–1588.

Andrews, M.J.I., Kontopidis, G., McInnes, C., Plater, A., Innes, L., Cowan, A., Jewsbury, P., and Fischer, P.M. (2006). REPLACE: a strategy for iterative design of cyclin-binding groove inhibitors. *Chembiochem: a European Journal of Chemical Biology* *7*, 1909–1915.

Atayde, V.D., Tschudi, C., and Ullu, E. (2011). The emerging world of small silencing RNAs in protozoan parasites. *Trends in Parasitology* *27*, 321–327.

Atayde, V.D., Ullu, E., and Kolev, N.G. (2012). A single-cloning-step procedure for the generation of RNAi plasmids producing long stem-loop RNA. *Molecular and Biochemical Parasitology* 184, 55–58.

Bahia, D., Oliveira, L.M., Lima, F.M., Oliveira, P., Da Silveira, J.F., Mortara, R.A., Ruiz, J.C., and Silveira, J.F. Da (2009). The TryPIKinome of five human pathogenic trypanosomatids: *Trypanosoma brucei*, *Trypanosoma cruzi*, *Leishmania major*, *Leishmania braziliensis* and *Leishmania infantum* - New tools for designing specific inhibitors. *Biochemical and Biophysical Research Communications* 390, 963–970.

Baker, N., Alsford, S., and Horn, D. (2011). Genome-wide RNAi screens in African trypanosomes identify the nifurtimox activator NTR and the eflornithine transporter AAT6. *Molecular and Biochemical Parasitology* 176, 55–57.

Bangs, J.D. (2011). Replication of the ERES:Golgi junction in bloodstream-form African trypanosomes. *Molecular Microbiology* 82, 1433–1443.

Barnes, R., Shi, H., Kolev, N., Tschudi, C., and Ullu, E. (2012). Comparative Genomics Reveals Two Novel RNAi Factors in *Trypanosoma brucei* and Provides Insight into the Core Machinery. *PLoS Pathogens* 8, e1002678.

Barouch-Bentov, R., and Sauer, K. (2011). Mechanisms of drug resistance in kinases. *Expert Opinion on Investigational Drugs* 20, 153–208.

Barquilla, A., Crespo, J.L., and Navarro, M. (2008). Rapamycin inhibits trypanosome cell growth by preventing TOR complex 2 formation. *Proceedings of the National Academy of Sciences of the United States of America* 105, 14579–14584.

Barquilla, A., Saldivia, M., Diaz, R., Bart, J.-M., Vidal, I., Calvo, E., Hall, M.N., and Navarro, M. (2012). Third target of rapamycin complex negatively regulates development of quiescence in *Trypanosoma brucei*. *Proceedings of the National Academy of Sciences of the United States of America*.

Barry, J.D., Hall, J.P.J., and Plenderleith, L. (2012). Genome hyperevolution and the success of a parasite. *Annals of the New York Academy of Sciences* 1267, 11–17.

Barry, J.D., Marcello, L., Morrison, L.J., Read, a F., Lythgoe, K., Jones, N., Carrington, M., Blandin, G., Böhme, U., Caler, E., et al. (2005). What the genome sequence is revealing about trypanosome antigenic variation. *Biochemical Society Transactions* 33, 986–989.

Belham, C., Roig, J., Caldwell, J.A., Aoyama, Y., Kemp, B.E., Comb, M., and Avruch, J. (2003). A mitotic cascade of NIMA family kinases. *Nercc1/Nek9* activates the *Nek6* and *Nek7* kinases. *The Journal of Biological Chemistry* 278, 34897–34909.

Bengs, F., Scholz, A., Kuhn, D., and Wiese, M. (2005). *LmxMPK9*, a mitogen-activated protein kinase homologue affects flagellar length in *Leishmania mexicana*. *Molecular Microbiology* 55, 1606–1615.

- Benz, C., Clucas, C., Mottram, J.C., and Hammarton, T.C. (2012). Cytokinesis in bloodstream stage *Trypanosoma brucei* requires a family of katanins and spastin. *PLoS One* 7, e30367.
- Benz, C., Mulindwa, J., Ouna, B., and Clayton, C. (2011). The *Trypanosoma brucei* zinc finger protein ZC3H18 is involved in differentiation. *Molecular and Biochemical Parasitology* 2–5.
- Berriman, M., Ghedin, E., Hertz-Fowler, C., Blandin, G., Renauld, H., Bartholomeu, D.C., Lennard, N.J., Caler, E., Hamlin, N.E., Haas, B., et al. (2005). The genome of the African trypanosome *Trypanosoma brucei*. *Science (New York, N.Y.)* 309, 416–422.
- Biebinger, S., Rettenmaier, S., Flaspohler, J., Hartmann, C., Peña-Díaz, J., Wirtz, L.E., Hotz, H.R., Barry, J.D., and Clayton, C. (1996). The PARP promoter of *Trypanosoma brucei* is developmentally regulated in a chromosomal context. *Nucleic Acids Research* 24, 1202–1211.
- Biebinger, S., Wirtz, L.E., Lorenz, P., and Clayton, C. (1997). Vectors for inducible expression of toxic gene products in bloodstream and procyclic *Trypanosoma brucei*. *Molecular and Biochemical Parasitology* 85, 99–112.
- Bishop, a, Buzko, O., Heyeck-Dumas, S., Jung, I., Kraybill, B., Liu, Y., Shah, K., Ulrich, S., Witucki, L., Yang, F., et al. (2000a). Unnatural ligands for engineered proteins: new tools for chemical genetics. *Annual Review of Biophysics and Biomolecular Structure* 29, 577–606.
- Bishop, a C., Ubersax, J. a, Petsch, D.T., Matheos, D.P., Gray, N.S., Blethrow, J., Shimizu, E., Tsien, J.Z., Schultz, P.G., Rose, M.D., et al. (2000b). A chemical switch for inhibitor-sensitive alleles of any protein kinase. *Nature* 407, 395–401.
- Bishop, A.C., and Shokat, K.M. (1999). Acquisition of inhibitor-sensitive protein kinases through protein design. *Pharmacology & Therapeutics* 82, 337–346.
- Boran, A.D.W., and Iyengar, R. (2010). Systems approaches to polypharmacology and drug discovery. *Current Opinion in Drug Discovery & Development* 13, 297–309.
- Bradley, B.A., and Quarmby, L.M. (2005). A NIMA-related kinase, Cnk2p, regulates both flagellar length and cell size in *Chlamydomonas*. *Journal of Cell Science* 118, 3317–3326.
- Brenchley, R., Tariq, H., McElhinney, H., Szöör, B., Huxley-Jones, J., Stevens, R., Matthews, K., and Taberner, L. (2007). The TriTryp phosphatome: analysis of the protein phosphatase catalytic domains. *BMC Genomics* 8, 434.
- Brennan, D.F., Dar, A.C., Hertz, N.T., Chao, W.C.H., Burlingame, A.L., Shokat, K.M., and Barford, D. (2011). A Raf-induced allosteric transition of KSR stimulates phosphorylation of MEK. *Nature* 472, 366–369.
- Brennand, A., Gualdrón-López, M., Coppens, I., Rigden, D.J., Ginger, M.L., and Michels, P. a M. (2011). Autophagy in parasitic protists: unique features and drug targets. *Molecular and Biochemical Parasitology* 177, 83–99.

- Brognard, J., and Hunter, T. (2011). Protein kinase signaling networks in cancer. *Current Opinion in Genetics & Development* 21, 4–11.
- Brown, M.D., and Sacks, D.B. (2009). Protein scaffolds in MAP kinase signalling. *Cellular Signalling* 21, 462–469.
- Brumlik, M.J., Pandeswara, S., Ludwig, S.M., Murthy, K., and Curiel, T.J. (2011). Parasite mitogen-activated protein kinases as drug discovery targets to treat human protozoan pathogens. *Journal of Signal Transduction* 2011, 971968.
- Brun, R., Blum, J., Chappuis, F., and Burri, C. (2010). Human African trypanosomiasis. *Lancet* 375, 148–159.
- Brun, R., Hecker, H., and Lun, Z.R. (1998). *Trypanosoma evansi* and *T. equiperdum*: distribution, biology, treatment and phylogenetic relationship (a review). *Veterinary Parasitology* 79, 95–107.
- Burleigh, B.A., Caler, E. V, Webster, P., and Andrews, N.W. (1997). A cytosolic serine endopeptidase from *Trypanosoma cruzi* is required for the generation of Ca²⁺ signaling in mammalian cells. *The Journal of Cell Biology* 136, 609–620.
- Burnett, G., and Kennedy, E.P. (1954). The enzymatic phosphorylation of proteins. *The Journal of Biological Chemistry* 211, 969–980.
- Burri, C., and Brun, R. (2003). Eflornithine for the treatment of human African trypanosomiasis. *Parasitology Research* 90 Supp 1, S49–52.
- Böhringer, S., and Hecker, H. (1974). Quantitative ultrastructural differences between strains of the *Trypanosoma brucei* subgroup during transformation in blood. *The Journal of Protozoology* 21, 694–698.
- Bütikofer, P., Ruepp, S., Boschung, M., and Roditi, I. (1997). “GPEET” procyclin is the major surface protein of procyclic culture forms of *Trypanosoma brucei brucei* strain 427. *The Biochemical Journal* 326 (Pt 2, 415–423.
- Bütikofer, P., Vassella, E., Ruepp, S., Boschung, M., Civenni, G., Seebeck, T., Hemphill, a, Mookherjee, N., Pearson, T.W., and Roditi, I. (1999). Phosphorylation of a major GPI-anchored surface protein of *Trypanosoma brucei* during transport to the plasma membrane. *Journal of Cell Science* 112 (Pt 1, 1785–1795.
- Chang, J., Baloh, R.H., and Milbrandt, J. (2009). The NIMA-family kinase Nek3 regulates microtubule acetylation in neurons. *Journal of Cell Science* 122, 2274–2282.
- Chaudhuri, M., Ott, R.D., and Hill, G.C. (2006). Trypanosome alternative oxidase: from molecule to function. *Trends in Parasitology* 22, 484–491.
- Chi, Y., and Clurman, B.E. (2010). Mass spectrometry-based identification of protein kinase substrates utilizing engineered kinases and thiophosphate labeling. *Current Protocols in Chemical Biology* 2, 1–19.
- Clapéron, A., and Therrien, M. (2007). KSR and CNK: two scaffolds regulating RAS-mediated RAF activation. *Oncogene* 26, 3143–3158.

Cleghorn, L. a T., Woodland, A., Collie, I.T., Torrie, L.S., Norcross, N., Luksch, T., Mpamhanga, C., Walker, R.G., Mottram, J.C., Brenk, R., et al. (2011). Identification of inhibitors of the *Leishmania* cdc2-related protein kinase CRK3. *ChemMedChem* 6, 2214–2224.

Cohen, P. (2002). The origins of protein phosphorylation. *Nature Cell Biology* 4, E127–30.

Coley, A.F., Dodson, H.C., Morris, M.T., and Morris, J.C. (2011). Glycolysis in the african trypanosome: targeting enzymes and their subcellular compartments for therapeutic development. *Molecular Biology International* 2011, 123702.

Cong, F., Cheung, A.K., and Huang, S.-M.A. (2012). Chemical genetics-based target identification in drug discovery. *Annual Review of Pharmacology and Toxicology* 52, 57–78.

Conrad, C., and Gerlich, D.W. (2010). Automated microscopy for high-content RNAi screening. *The Journal of Cell Biology* 188, 453–461.

Cuny, G.D. (2009). Kinase inhibitors as potential therapeutics for acute and chronic neurodegenerative conditions. *Current Pharmaceutical Design* 15, 3919–3939.

Czichos, J., Nonnengaesser, C., and Overath, P. (1986). *Trypanosoma brucei*: cis-aconitate and temperature reduction as triggers of synchronous transformation of bloodstream to procyclic trypomastigotes in vitro. *Experimental Parasitology* 62, 283–291.

Dean, S., Marchetti, R., Kirk, K., and Matthews, K.R. (2009). A surface transporter family conveys the trypanosome differentiation signal. *Nature* 459, 213–217.

DeGrasse, J. a, DuBois, K.N., Devos, D., Siegel, T.N., Sali, A., Field, M.C., Rout, M.P., and Chait, B.T. (2009). Evidence for a shared nuclear pore complex architecture that is conserved from the last common eukaryotic ancestor. *Molecular & Cellular Proteomics : MCP* 8, 2119–2130.

Deininger, M.W., and Druker, B.J. (2003). Specific targeted therapy of chronic myelogenous leukemia with imatinib. 55, 401–423.

Diaz-Gonzalez, R., Kuhlmann, F.M., Galan-Rodriguez, C., Madeira da Silva, L., Saldivia, M., Karver, C.E., Rodriguez, A., Beverley, S.M., Navarro, M., and Pollastri, M.P. (2011). The Susceptibility of Trypanosomatid Pathogens to PI3/mTOR Kinase Inhibitors Affords a New Opportunity for Drug Repurposing. *PLoS Neglected Tropical Diseases* 5, e1297.

Doerig, C., Billker, O., Pratt, D., and Endicott, J. (2005). Protein kinases as targets for antimalarial intervention: Kinomics, structure-based design, transmission-blockade, and targeting host cell enzymes. *Biochimica Et Biophysica Acta* 1754, 132–150.

Domenicali Pfister, D., Burkard, G., Morand, S., Renggli, C.K., Roditi, I., and Vassella, E. (2006a). A Mitogen-activated protein kinase controls differentiation of bloodstream forms of *Trypanosoma brucei*. *Eukaryotic Cell* 5, 1126–1135.

Domenicali Pfister, D., Burkard, G., Morand, S., Renggli, C.K., Roditi, I., and Vassella, E. (2006b). A Mitogen-activated protein kinase controls differentiation of bloodstream forms of *Trypanosoma brucei*. *Eukaryotic Cell* 5, 1126–1135.

Dorin, D., Le Roch, K., Sallicandro, P., Alano, P., Parzy, D., Poulet, P., Meijer, L., and Doerig, C. (2001). Pfnek-1, a NIMA-related kinase from the human malaria parasite *Plasmodium falciparum* Biochemical properties and possible involvement in MAPK regulation. *European Journal of Biochemistry / FEBS* 268, 2600–2608.

Dorin-Semblat, D., Schmitt, S., Semblat, J.-P., Sicard, A., Reininger, L., Goldring, D., Patterson, S., Quashie, N., Chakrabarti, D., Meijer, L., et al. (2011). *Plasmodium falciparum* NIMA-related kinase Pfnek-1: sex specificity and assessment of essentiality for the erythrocytic asexual cycle. *Microbiology (Reading, England)* 157, 2785–2794.

Drew, M.E., Morris, J.C., Wang, Z., Wells, L., Sanchez, M., Landfear, S.M., and Englund, P.T. (2003). The adenosine analog tubercidin inhibits glycolysis in *Trypanosoma brucei* as revealed by an RNA interference library. *The Journal of Biological Chemistry* 278, 46596–46600.

Druker, B.J. (2004). Imatinib as a paradigm of targeted therapies. *Advances in Cancer Research* 91, 1–30.

Dumon-Seignovert, L., Cariot, G., and Vuillard, L. (2004). The toxicity of recombinant proteins in *Escherichia coli*: a comparison of overexpression in BL21(DE3), C41(DE3), and C43(DE3). *Protein Expression and Purification* 37, 203–206.

Durand-Dubief, M., Kohl, L., and Bastin, P. (2003). Efficiency and specificity of RNA interference generated by intra- and intermolecular double stranded RNA in *Trypanosoma brucei*. *Molecular and Biochemical Parasitology* 129, 11–21.

Duszenko, M., Ginger, M.L., Brennand, A., Gualdrón-López, M., Colombo, M.I., Coombs, G.H., Coppens, I., Jayabalasingham, B., Langsley, G., Lisboa de Castro, S., et al. (2011). Autophagy in protists. *Autophagy* 7, 127–158.

D'Elia, M. a, Pereira, M.P., and Brown, E.D. (2009). Are essential genes really essential? *Trends in Microbiology* 17, 433–438.

Ellis, J., Sarkar, M., Hendriks, E., and Matthews, K. (2004). A novel ERK-like, CRK-like protein kinase that modulates growth in *Trypanosoma brucei* via an autoregulatory C-terminal extension. *Molecular Microbiology* 53, 1487–1499.

Elphick, L.M., Lee, S.E., Gouverneur, V., and Mann, D.J. (2007). Using chemical genetics and ATP analogues to dissect protein kinase function. *ACS Chemical Biology* 2, 299–314.

Emmer, B.T., Nakayasu, E.S., Souther, C., Choi, H., Sobreira, T.J.P., Epting, C.L., Nesvizhskii, A.I., Almeida, I.C., and Engman, D.M. (2011). Global analysis of protein palmitoylation in African trypanosomes. *Eukaryotic Cell* 10, 455–463.

Engstler, M., and Boshart, M. (2004). Cold shock and regulation of surface protein trafficking convey sensitization to inducers of stage differentiation in *Trypanosoma brucei*. *Genes & Development* 18, 2798–2811.

Engstler, M., Pfohl, T., Herminghaus, S., Boshart, M., Wiegertjes, G., Heddergott, N., and Overath, P. (2007). Hydrodynamic flow-mediated protein sorting on the cell surface of trypanosomes. *Cell* 131, 505–515.

Erdmann, M., Scholz, A., Melzer, I.M., Schmetz, C., and Wiese, M. (2006). Interacting protein kinases involved in the regulation of flagellar length. *Molecular Biology of the Cell* 17, 2035–2045.

Farr, H., and Gull, K. (2012). Cytokinesis in trypanosomes. *Cytoskeleton* (Hoboken, N.J.) 000,.

Le Febvre, R.B., and Hill, G.C. (1986). Effects on *Trypanosoma brucei* differentiating bloodstream trypomastigotes and established procyclic trypomastigotes when grown in the presence of respiratory inhibitors. *The Journal of Parasitology* 72, 481–483.

Fedorov, O., Müller, S., and Knapp, S. (2010). The (un)targeted cancer kinome. *Nature Chemical Biology* 6, 166–169.

Ferrante, A., and Allison, A.C. (1983). Alternative pathway activation of complement by African trypanosomes lacking a glycoprotein coat. *Parasite Immunology* 5, 491–498.

Field, M.C., Allen, C.L., Dhir, V., Goulding, D., Hall, B.S., Morgan, G.W., Veazey, P., and Engstler, M. (2004). New approaches to the microscopic imaging of *Trypanosoma brucei*. *Microscopy and Microanalysis : the Official Journal of Microscopy Society of America, Microbeam Analysis Society, Microscopical Society of Canada* 10, 621–636.

Fire, A., Xu, S.Q., Montgomery, M.K.K., Kostas, S.A.A., Driver, S.E.E., and Mello, C.C.C. (1998). Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 391, 806–811.

Fischer, E.H., Graves, D.J., Crittenden, E.R.S., and Krebs, E.G. (1959). Structure of the site phosphorylated in the phosphorylase b to a reaction. *The Journal of Biological Chemistry* 234, 1698–1704.

Fischer, E.H., and Krebs, E.G. (1955). Conversion of phosphorylase b to phosphorylase a in muscle extracts. *The Journal of Biological Chemistry* 216, 121–132.

Flaspohler, J. a, Jensen, B.C., Saveria, T., Kifer, C.T., and Parsons, M. (2010). A novel protein kinase localized to lipid droplets is required for droplet biogenesis in trypanosomes. *Eukaryotic Cell* 9, 1702–1710.

Fragoso, C.M., Schumann Burkard, G., Oberle, M., Renggli, C.K., Hilzinger, K., and Roditi, I. (2009). PSSA-2, a membrane-spanning phosphoprotein of *Trypanosoma brucei*, is required for efficient maturation of infection. *PloS One* 4, e7074.

Frearson, J.A., Wyatt, P.G., Gilbert, I.H., and Fairlamb, A.H. (2007). Target assessment for antiparasitic drug discovery. *23*, 589–595.

Fu, C., Wehr, D.R., Edwards, J., and Hauge, B. (2008). Rapid one-step recombinational cloning. *Nucleic Acids Research* *36*, e54.

Gale, M., and Parsons, M. (1993). A *Trypanosoma brucei* gene family encoding protein kinases with catalytic domains structurally related to Nek1 and NIMA. *Molecular and Biochemical Parasitology* *59*, 111–121.

Gamsjaeger, R., Liew, C.K., Loughlin, F.E., Crossley, M., and Mackay, J.P. (2007). Sticky fingers: zinc-fingers as protein-recognition motifs. *Trends in Biochemical Sciences* *32*, 63–70.

Garske, A.L., Peters, U., Cortesi, A.T., Perez, J.L., and Shokat, K.M. (2011). Chemical genetic strategy for targeting protein kinases based on covalent complementarity. *Proceedings of the National Academy of Sciences of the United States of America* *108*, 15046–15052.

Ghoreschi, K., Laurence, A., and O'Shea, J.J. (2009). Selectivity and therapeutic inhibition of kinases: to be or not to be? *Nature Immunology* *10*, 356–360.

Giaever, G., Chu, A.M., Ni, L., Connelly, C., Riles, L., Véronneau, S., Dow, S., Lucau-Danila, A., Anderson, K., André, B., et al. (2002). Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* *418*, 387–391.

Glover, L., and Horn, D. (2009). Site-specific DNA double-strand breaks greatly increase stable transformation efficiency in *Trypanosoma brucei*. *Molecular and Biochemical Parasitology* *166*, 194–197.

Gomes, F.C., Ali, N.O.M., Brown, E., Walker, R.G., Grant, K.M., and Mottram, J.C. (2010). Recombinant *Leishmania mexicana* CRK3:CYCA has protein kinase activity in the absence of phosphorylation on the T-loop residue Thr178. *Molecular and Biochemical Parasitology* *171*, 89–96.

Good, M.C., Zalatan, J.G., and Lim, W. a. (2011). Scaffold Proteins: Hubs for Controlling the Flow of Cellular Information. *Science* *332*, 680–686.

De Graffenried, C.L., Ho, H.H., and Warren, G. (2008). Polo-like kinase is required for Golgi and bilobe biogenesis in *Trypanosoma brucei*. *181*, 431–438.

Grant, K.M., Dunion, M.H., Yardley, V., Skaltsounis, A., Marko, D., Eisenbrand, G., Croft, S.L., Meijer, L., and Mottram, J.C. (2004). Inhibitors of *Leishmania mexicana* CRK3 cyclin-dependent kinase: chemical library screen and antileishmanial activity. *Antimicrobial Agents and Chemotherapy* *48*, 3033–3042.

Gualdrón-López, M., Brennand, A., Hannaert, V., Quiñones, W., Cáceres, A.J., Bringaud, F., Concepción, J.L., and Michels, P. a M. (2012). When, how and why glycolysis became compartmentalised in the Kinetoplastea. A new look at an ancient organelle. *International Journal for Parasitology* *42*, 1–20.

Gunasekera, K., Wüthrich, D., Braga-Lagache, S., Heller, M., and Ochsenreiter, T. (2012). Proteome remodelling during development from blood to insect-form

- Trypanosoma brucei quantified by SILAC and mass spectrometry. *BMC Genomics* 13, 556.
- Günzl, A. (2010). The pre-mRNA splicing machinery of trypanosomes: complex or simplified? *Eukaryotic Cell* 9, 1159–1170.
- Güttinger, A., Schwab, C., Morand, S., Roditi, I., and Vassella, E. (2007). A mitogen-activated protein kinase of *Trypanosoma brucei* confers resistance to temperature stress. *Molecular and Biochemical Parasitology* 153, 203–206.
- Haanstra, J.R., Kerkhoven, E.J., Van Tuijl, A., Blits, M., Wurst, M., Van Nuland, R., Albert, M.-A., Michels, P. a M., Bouwman, J., Clayton, C., et al. (2011). A domino effect in drug action: from metabolic assault towards parasite differentiation. *Molecular Microbiology* 79, 94–108.
- Hammarström, M., Woestenenk, E. a, Hellgren, N., Härd, T., and Berglund, H. (2006). Effect of N-terminal solubility enhancing fusion proteins on yield of purified target protein. *Journal of Structural and Functional Genomics* 7, 1–14.
- Hammarton, T.C., Kramer, S., Tetley, L., Boshart, M., and Mottram, J.C. (2007). *Trypanosoma brucei* Polo-like kinase is essential for basal body duplication, kDNA segregation and cytokinesis. *Molecular Microbiology* 65, 1229–1248.
- Hanks, S.K. (2003). Genomic analysis of the eukaryotic protein kinase superfamily: a perspective. *Genome Biology* 4, 111.
- Hanks, S.K., and Hunter, T. (1995). Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *FASEB Journal : Official Publication of the Federation of American Societies for Experimental Biology* 9, 576–596.
- Hanks, S.K., Quinn, a M., and Hunter, T. (1988). The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. *Science (New York, N.Y.)* 241, 42–52.
- Hann, M.M., and Oprea, T.I. (2004). Pursuing the leadlikeness concept in pharmaceutical research. *Current Opinion in Chemical Biology* 8, 255–263.
- Hassan, P., Fergusson, D., Grant, K.M., and Mottram, J.C. (2001). The CRK3 protein kinase is essential for cell cycle progression of *Leishmania mexicana*. *Molecular and Biochemical Parasitology* 113, 189–198.
- Von Heijne, G. (2006). Membrane-protein topology. *Nature Reviews. Molecular Cell Biology* 7, 909–918.
- Helliwell, C.A., Wesley, S. V, Wielopolska, A.J., and Waterhouse, P.M. (2002). High-throughput vectors for efficient gene silencing in plants. *Functional Plant Biology* 29, 1217–1225.
- Helms, M.J., Ambit, A., Appleton, P., Tetley, L., Coombs, G.H., and Mottram, J.C. (2006). Bloodstream form *Trypanosoma brucei* depend upon multiple metacaspases associated with RAB11-positive endosomes. *Journal of Cell Science* 119, 1105–1117.

- Herm-Götz, A., Agop-Nersesian, C., Münter, S., Grimley, J.S., Wandless, T.J., Frischknecht, F., and Meissner, M. (2007). Rapid control of protein level in the apicomplexan *Toxoplasma gondii*. *Nature Methods* 4, 1003–1005.
- Hirumi, H. Hirumi, K. (1989). Continuous cultivation of *Trypanosoma brucei* blood stream forms in a medium containing a low concentration of serum protein without feeder cell layers. *The Journal of Parasitology* 75, 985–989.
- Hopkins, A.L. (2008). Network pharmacology: the next paradigm in drug discovery. *Nature Chemical Biology* 4, 682–690.
- Howell, J.J., and Manning, B.D. (2011). mTOR couples cellular nutrient sensing to organismal metabolic homeostasis. *Trends in Endocrinology and Metabolism: TEM* 22, 94–102.
- Hua, S.B., and Wang, C.C. (1994). Differential accumulation of a protein kinase homolog in *Trypanosoma brucei*. *Journal of Cellular Biochemistry* 54, 20–31.
- Hua, S.B., and Wang, C.C. (1997). Interferon-gamma activation of a mitogen-activated protein kinase, KFR1, in the bloodstream form of *Trypanosoma brucei*. *The Journal of Biological Chemistry* 272, 10797–10803.
- Ikeda, K. (2012). Polo-like kinase is necessary for flagellum inheritance in *Trypanosoma brucei*. *Journal of Cell Science*.
- Ivens, A.C., Peacock, C.S., Worthey, E.A., Murphy, L., Aggarwal, G., Berriman, M., Sisk, E., Rajandream, M.-A., Adlem, E., Aert, R., et al. (2005). The genome of the kinetoplastid parasite, *Leishmania major*. *Science (New York, N.Y.)* 309, 436–442.
- Jackson, A.P. (2007). Evolutionary consequences of a large duplication event in *Trypanosoma brucei*: chromosomes 4 and 8 are partial duplicons. *BMC Genomics* 8, 432.
- Jackson, D.G., Owen, M.J., and Voorheis, H.P. (1985). A new method for the rapid purification of both the membrane-bound and released forms of the variant surface glycoprotein from *Trypanosoma brucei*. *The Biochemical Journal* 230, 195–202.
- Jamonneau, V., Ilboudo, H., Kaboré, J., Kaba, D., Koffi, M., Solano, P., Garcia, A., Courtin, D., Laveissière, C., Lingue, K., et al. (2012). Untreated Human Infections by *Trypanosoma brucei gambiense* Are Not 100% Fatal. *PLoS Neglected Tropical Diseases* 6, e1691.
- Jensen, B.C., Kifer, C.T., and Parsons, M. (2011). *Trypanosoma brucei*: Two mitogen activated protein kinase kinases are dispensable for growth and virulence of the bloodstream form. *Experimental Parasitology* 1–6.
- Jensen, B.C., Sivam, D., Kifer, C.T., Myler, P.J., and Parsons, M. (2009). Widespread variation in transcript abundance within and across developmental stages of *Trypanosoma brucei*. *BMC Genomics* 10, 482.
- Jetton, N., Rothberg, K.G., Hubbard, J.G., Wise, J., Li, Y., Ball, H.L., and Ruben, L. (2009). The cell cycle as a therapeutic target against *Trypanosoma brucei*:

Hesperadin inhibits Aurora kinase-1 and blocks mitotic progression in bloodstream forms. *Molecular Microbiology* 72, 442–458.

Johnson, S.A., and Hunter, T. (2005). Kinomics: methods for deciphering the kinome. *Nature Methods* 2, 17–25.

Kabani, S., Fenn, K., Ross, A., Ivens, A., Smith, T.K., Ghazal, P., and Matthews, K. (2009). Genome-wide expression profiling of in vivo-derived bloodstream parasite stages and dynamic analysis of mRNA alterations during synchronous differentiation in *Trypanosoma brucei*. *BMC Genomics* 10, 427.

Kalidas, S., Li, Q., and Phillips, M. a (2011). A Gateway® compatible vector for gene silencing in bloodstream form *Trypanosoma brucei*. *Molecular and Biochemical Parasitology* 178, 51–55.

Karaman, M.W., Herrgard, S., Treiber, D.K., Gallant, P., Atteridge, C.E., Campbell, B.T., Chan, K.W., Ciceri, P., Davis, M.I., Edeen, P.T., et al. (2008). A quantitative analysis of kinase inhibitor selectivity. *Nature Biotechnology* 26, 127–132.

Kelly, S., Reed, J., Kramer, S., Ellis, L., Webb, H., Sunter, J., Salje, J., Marinsek, N., Gull, K., Wickstead, B., et al. (2007). Functional genomics in *Trypanosoma brucei*: a collection of vectors for the expression of tagged proteins from endogenous and ectopic gene loci. *Molecular and Biochemical Parasitology* 154, 103–109.

Klug, a (1999). Zinc finger peptides for the regulation of gene expression. *Journal of Molecular Biology* 293, 215–218.

Knight, Z. a, Lin, H., and Shokat, K.M. (2010). Targeting the cancer kinome through polypharmacology. *Nature Reviews. Cancer* 10, 130–137.

Knight, Z. a, and Shokat, K.M. (2005). Features of selective kinase inhibitors. *Chemistry & Biology* 12, 621–637.

Knighton, D.R., Zheng, J.H., Ten Eyck, L.F., Ashford, V.A., Xuong, N.H., Taylor, S.S., and Sowadski, J.M. (1991). Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. *Science (New York, N.Y.)* 253, 407–414.

Kokuryo, T., Senga, T., Yokoyama, Y., Nagino, M., Nimura, Y., and Hamaguchi, M. (2007). Nek2 as an effective target for inhibition of tumorigenic growth and peritoneal dissemination of cholangiocarcinoma. *Cancer Research* 67, 9637–9642.

Kolev, N.G., Franklin, J.B., Carmi, S., Shi, H., Michaeli, S., and Tschudi, C. (2010). The Transcriptome of the Human Pathogen *Trypanosoma brucei* at Single-Nucleotide Resolution. *PLoS Pathogens* 6, e1001090.

Kolev, N.G., Tschudi, C., and Ullu, E. (2011). RNA interference in protozoan parasites: achievements and challenges. *Eukaryotic Cell* 10, 1156–1163.

Kontopidis, G., Andrews, M.J.I., McInnes, C., Cowan, A., Powers, H., Innes, L., Plater, A., Griffiths, G., Paterson, D., Zheleva, D.I., et al. (2003). Insights into

cyclin groove recognition: complex crystal structures and inhibitor design through ligand exchange. *Structure* (London, England : 1993) *11*, 1537–1546.

Korfel, A., and Thiel, E. (2007). Targeted therapy and blood-brain barrier. *Recent Results in Cancer Research. Fortschritte Der Krebsforschung. Progrès Dans Les Recherches Sur Le Cancer* *176*, 123–133.

Koumandou, V.L., Natesan, S.K. a, Sergeenko, T., and Field, M.C. (2008). The trypanosome transcriptome is remodelled during differentiation but displays limited responsiveness within life stages. *BMC Genomics* *9*, 298.

Kramer, S., Kimblin, N.C., and Carrington, M. (2010). Genome-wide in silico screen for CCCH-type zinc finger proteins of *Trypanosoma brucei*, *Trypanosoma cruzi* and *Leishmania major*. *BMC Genomics* *11*, 283.

Krebs, E.G., and Fischer, E.H. (1956). The phosphorylase b to a converting enzyme of rabbit skeletal muscle. *Biochimica Et Biophysica Acta* *20*, 150–157.

Kröncke, K.-D., and Klotz, L.-O. (2009). Zinc fingers as biologic redox switches? *Antioxidants & Redox Signaling* *11*, 1015–1027.

Kutateladze, T., and Overduin, M. (2001). Structural mechanism of endosome docking by the FYVE domain. *Science* (New York, N.Y.) *291*, 1793–1796.

Lambrecht, F.L. (1985). Trypanosomes and Hominid Evolution. *BioScience* *35*, 640–646.

Laurell, E., Beck, K., Krupina, K., Theerthagiri, G., Bodenmiller, B., Horvath, P., Aebersold, R., Antonin, W., and Kutay, U. (2011). Phosphorylation of Nup98 by multiple kinases is crucial for NPC disassembly during mitotic entry. *Cell* *144*, 539–550.

LaVallie, E.R., DiBlasio, E.A., Kovacic, S., Grant, K.L., Schendel, P.F., and McCoy, J.M. (1993). A thioredoxin gene fusion expression system that circumvents inclusion body formation in the *E. coli* cytoplasm. *Bio/technology* (Nature Publishing Company) *11*, 187–193.

Laxman, S., Riechers, A., Sadilek, M., Schwede, F., and Beavo, J. a (2006). Hydrolysis products of cAMP analogs cause transformation of *Trypanosoma brucei* from slender to stumpy-like forms. *Proceedings of the National Academy of Sciences of the United States of America* *103*, 19194–19199.

Lemmon, M. (2004). Pleckstrin homology domains: not just for phosphoinositides. *Biochemical Society Transactions* *32*, 707.

Lemmon, M. a, Ferguson, K.M., and Abrams, C.S. (2002). Pleckstrin homology domains and the cytoskeleton. *FEBS Letters* *513*, 71–76.

Li, Z. (2012). Regulation of the cell division cycle in *Trypanosoma brucei*. *Eukaryotic Cell*.

Li, Z., Gourguechon, S., and Wang, C.C. (2007a). Tousled-like kinase in a microbial eukaryote regulates spindle assembly and S-phase progression by

interacting with Aurora kinase and chromatin assembly factors. *Journal of Cell Science* 120, 3883–3894.

Li, Z., Gourguechon, S., and Wang, C.C. (2007b). Tousled-like kinase in a microbial eukaryote regulates spindle assembly and S-phase progression by interacting with Aurora kinase and chromatin assembly factors. *Journal of Cell Science* 120, 3883–3894.

Li, Z., Umeyama, T., and Wang, C.C. Polo-like kinase guides cytokinesis in *Trypanosoma brucei* through an indirect means. *Eukaryotic Cell* 9, 705–716.

Li, Z., Umeyama, T., and Wang, C.C. (2008). The chromosomal passenger complex and a mitotic kinesin interact with the Tousled-like kinase in trypanosomes to regulate mitosis and cytokinesis. *PLoS One* 3, e3814.

Li, Z., Umeyama, T., and Wang, C.C. (2009). The Aurora Kinase in *Trypanosoma brucei* plays distinctive roles in metaphase-anaphase transition and cytokinetic initiation. *PLoS Pathogens* 5, e1000575.

Li, Z., and Wang, C.C. (2006). Changing roles of aurora-B kinase in two life cycle stages of *Trypanosoma brucei*. *Eukaryotic Cell* 5, 1026–1035.

Liu, B.Y., Liu, Y.N., Motyka, S.A., Agbo, E.E.C., and Englund, P.T. (2005). Fellowship of the rings: the replication of kinetoplast DNA. *21*, 363–369.

Liu, Y., and Gray, N.S. (2006). Rational design of inhibitors that bind to inactive kinase conformations. *Nature Chemical Biology* 2, 358–364.

Lizcano, J.M., Deak, M., Morrice, N., Kieloch, A., Hastie, C.J., Dong, L., Schutkowski, M., Reimer, U., and Alessi, D.R. (2002). Molecular basis for the substrate specificity of NIMA-related kinase-6 (NEK6). Evidence that NEK6 does not phosphorylate the hydrophobic motif of ribosomal S6 protein kinase and serum- and glucocorticoid-induced protein kinase in vivo. *The Journal of Biological Chemistry* 277, 27839–27849.

Lourido, S., Shuman, J., Zhang, C., Shokat, K.M., Hui, R., and Sibley, L.D. (2010). Calcium-dependent protein kinase 1 is an essential regulator of exocytosis in *Toxoplasma*. *Nature* 465, 359–362.

Low, H., Chua, C.S., and Sim, T.-S. (2012). *Plasmodium falciparum* possesses a unique dual-specificity serine/threonine and tyrosine kinase, Pfnek3. *Cellular and Molecular Life Sciences : CMLS* 69, 1523–1535.

Low, H., Lye, Y.M., and Sim, T.-S. (2007). Pfnek3 functions as an atypical MAPKK in *Plasmodium falciparum*. *Biochemical and Biophysical Research Communications* 361, 439–444.

Lye, Y.M., Chan, M., and Sim, T.-S. (2006). Pfnek3: an atypical activator of a MAP kinase in *Plasmodium falciparum*. *FEBS Letters* 580, 6083–6092.

Ma, J., Benz, C., Grimaldi, R., Stockdale, C., Wyatt, P., Frearson, J., and Hammarton, T.C. (2010). Nuclear DBF-2-related kinases are essential regulators

of cytokinesis in bloodstream stage *Trypanosoma brucei*. *The Journal of Biological Chemistry* *285*, 15356–15368.

Mackey, Z.B., Koupparis, K., Nishino, M., and McKerrow, J.H. (2011). High-throughput analysis of an RNAi library identifies novel kinase targets in *Trypanosoma brucei*. *Chemical Biology & Drug Design* *78*, 454–463.

Mahjoub, M.R., Qasim Rasi, M., and Quarmby, L.M. (2004). A NIMA-related kinase, Fa2p, localizes to a novel site in the proximal cilia of *Chlamydomonas* and mouse kidney cells. *Molecular Biology of the Cell* *15*, 5172–5186.

Mahjoub, M.R., Trapp, M.L., and Quarmby, L.M. (2005). NIMA-related kinases defective in murine models of polycystic kidney diseases localize to primary cilia and centrosomes. *Journal of the American Society of Nephrology : JASN* *16*, 3485–3489.

Malvy, D., and Chappuis, F. (2011). Sleeping sickness. *Clinical Microbiology and Infection : the Official Publication of the European Society of Clinical Microbiology and Infectious Diseases* *17*, 986–995.

Manning, G., Plowman, G.D., Hunter, T., and Sudarsanam, S. (2002a). Evolution of protein kinase signaling from yeast to man. *Trends in Biochemical Sciences* *27*, 514–520.

Manning, G., Reiner, D.S., Lauwaet, T., Dacre, M., Smith, A., Zhai, Y., Svard, S., and Gillin, F.D. (2011). The minimal kinome of *Giardia lamblia* illuminates early kinase evolution and unique parasite biology. *Genome Biology* *12*, R66.

Manning, G., Whyte, D.B., Martinez, R., Hunter, T., and Sudarsanam, S. (2002b). The protein kinase complement of the human genome. *298*, 1912–+.

Matthews, K.R. (2005). The developmental cell biology of *Trypanosoma brucei*. *Journal of Cell Science* *118*, 283–290.

Matthews, K.R., Ellis, J.R., and Paterou, A. (2004). Molecular regulation of the life cycle of African trypanosomes. *Trends in Parasitology* *20*, 40–47.

Matthews, K.R., and Gull, K. (1994). Evidence for an interplay between cell cycle progression and the initiation of differentiation between life cycle forms of African trypanosomes. *The Journal of Cell Biology* *125*, 1147–1156.

May, S.F., Peacock, L., Almeida Costa, C.I.C., Gibson, W.C., Tetley, L., Robinson, D.R., and Hammarton, T.C. (2012). The *Trypanosoma brucei* AIR9-like protein is cytoskeleton-associated and is required for nucleus positioning and accurate cleavage furrow placement. *Molecular Microbiology*.

Mercer, L., Bowling, T., Perales, J., Freeman, J., Nguyen, T., Bacchi, C., Yarlett, N., Don, R., Jacobs, R., and Nare, B. (2011). 2,4-Diaminopyrimidines as Potent Inhibitors of *Trypanosoma brucei* and Identification of Molecular Targets by a Chemical Proteomics Approach. *PLoS Neglected Tropical Diseases* *5*, e956.

Michels, P. a M., Bringaud, F., Herman, M., and Hannaert, V. (2006). Metabolic functions of glycosomes in trypanosomatids. *Biochimica Et Biophysica Acta* 1763, 1463–1477.

Moniz, L., Dutt, P., Haider, N., and Stambolic, V. (2011). Nek family of kinases in cell cycle, checkpoint control and cancer. *Cell Division* 6, 18.

Monnerat, S., Clucas, C., Brown, E., Mottram, J.C., and Hammarton, T.C. (2009). Searching for novel cell cycle regulators in *Trypanosoma brucei* with an RNA interference screen. *BMC Research Notes* 2, 46.

Mora, A., Komander, D., Van Aalten, D.M.F., and Alessi, D.R. (2004). PDK1, the master regulator of AGC kinase signal transduction. *Seminars in Cell & Developmental Biology* 15, 161–170.

Morales, M. a, Watanabe, R., Dacher, M., Chafey, P., Osorio y Fortéa, J., Scott, D. a, Beverley, S.M., Ommen, G., Clos, J., Hem, S., et al. (2010). Phosphoproteome dynamics reveal heat-shock protein complexes specific to the *Leishmania donovani* infectious stage. *Proceedings of the National Academy of Sciences of the United States of America* 107, 8381–8386.

Morand, S., Renggli, C.K., Roditi, I., and Vassella, E. (2012). MAP kinase kinase 1 (MKK1) is essential for transmission of *Trypanosoma brucei* by *Glossina morsitans*. *Molecular and Biochemical Parasitology* 1, 1–4.

Morris, J.C., Wang, Z., Drew, M.E., and Englund, P.T. (2002). Glycolysis modulates trypanosome glycoprotein expression as revealed by an RNAi library. *The EMBO Journal* 21, 4429–4438.

Morty, R.E., Lonsdale-Eccles, J.D., Mentele, R., Auerswald, E.A., and Coetzer, T.H. (2001). Trypanosome-derived oligopeptidase B is released into the plasma of infected rodents, where it persists and retains full catalytic activity. *Infection and Immunity* 69, 2757–2761.

Mowatt, M.R., and Clayton, C.E. (1987). Developmental regulation of a novel repetitive protein of *Trypanosoma brucei*. *Molecular and Cellular Biology* 7, 2838–2844.

Mowatt, M.R., Wisdom, G.S., and Clayton, C.E. (1989). Variation of tandem repeats in the developmentally regulated procyclic acidic repetitive proteins of *Trypanosoma brucei*. *Molecular and Cellular Biology* 9, 1332–1335.

Munday, J.C., McLuskey, K., Brown, E., Coombs, G.H., and Mottram, J.C. (2011). Oligopeptidase B deficient mutants of *Leishmania major*. *Molecular and Biochemical Parasitology* 175, 49–57.

Mäser, P., Wittlin, S., Rottmann, M., Wenzler, T., Kaiser, M., and Brun, R. (2012). Antiparasitic agents: new drugs on the horizon. *Current Opinion in Pharmacology* 1–5.

Müller, I.B., Domenicali-Pfister, D., Roditi, I., and Vassella, E. (2002a). Stage-specific requirement of a mitogen-activated protein kinase by *Trypanosoma brucei*. *Molecular Biology of the Cell* 13, 3787–3799.

- Müller, I.B., Domenicali-Pfister, D., Roditi, I., and Vassella, E. (2002b). Stage-specific requirement of a mitogen-activated protein kinase by *Trypanosoma brucei*. *Molecular Biology of the Cell* 13, 3787–3799.
- Nakano, H., Kobayashi, E., Takahashi, I., Tamaoki, T., Kuzuu, Y., and Iba, H. (1987). Staurosporine inhibits tyrosine-specific protein kinase activity of Rous sarcoma virus transforming protein p60. *The Journal of Antibiotics* 40, 706–708.
- Naula, C., Parsons, M., and Mottram, J.C. (2005). Protein kinases as drug targets in trypanosomes and *Leishmania*. *Biochimica Et Biophysica Acta* 1754, 151–159.
- Nett, I.R.E., Davidson, L., Lamont, D., and Ferguson, M.A.J. (2009a). Identification and specific localization of tyrosine-phosphorylated proteins in *Trypanosoma brucei*. *Eukaryotic Cell* 8, 617–626.
- Nett, I.R.E., Martin, D.M.A., Miranda-Saavedra, D., Lamont, D., Barber, J.D., Mehlert, A., and Ferguson, M.A.J. (2009b). The phosphoproteome of bloodstream form *Trypanosoma brucei*, causative agent of African sleeping sickness. *Molecular & Cellular Proteomics : MCP* 8, 1527–1538.
- Neumann, T., Junker, H.-D., Schmidt, K., and Sekul, R. (2007). SPR-based fragment screening: advantages and applications. *Current Topics in Medicinal Chemistry* 7, 1630–1642.
- Ngô, H., Tschudi, C., Gull, K., and Ullu, E. (1998). Double-stranded RNA induces mRNA degradation in *Trypanosoma brucei*. *Proceedings of the National Academy of Sciences of the United States of America* 95, 14687–14692.
- Novy, R., and Yaeger, K. (1996). Ligation independent cloning: efficient directional cloning of PCR products. *InNovations* 6–8.
- Oberholzer, M., Langousis, G., Nguyen, H.T., Saada, E. a, Shimogawa, M.M., Jonsson, Z.O., Nguyen, S.M., Wohlschlelgel, J. a, and Hill, K.L. (2011). Independent analysis of the flagellum surface and matrix proteomes provides insight into flagellum signaling in mammalian-infectious *Trypanosoma brucei*. *Molecular & Cellular Proteomics : MCP*.
- Ochiana, S.O., Pandarinath, V., Wang, Z., Kapoor, R., Ondrechen, M.J., Ruben, L., and Pollastri, M.P. (2012). The human Aurora kinase inhibitor danusertib is a lead compound for anti-trypanosomal drug discovery via target repurposing. *European Journal of Medicinal Chemistry*.
- Ojo, K.K., Arakaki, T.L., Napuli, A.J., Inampudi, K.K., Keyloun, K.R., Zhang, L., Hol, W.G.J., Verlinde, C.L.M.J., Merritt, E. a, and Van Voorhis, W.C. (2010). Structure determination of glycogen synthase kinase-3 from *Leishmania major* and comparative inhibitor structure-activity relationships with *Trypanosoma brucei* GSK-3. *Molecular and Biochemical Parasitology* 1–11.
- Ojo, K.K., Gillespie, J.R., Riechers, A.J., Napuli, A.J., Verlinde, C.L.M.J.C., Buckner, F.S., Gelb, M.H., Domostoj, M.M., Wells, S.J., Scheer, A., et al. (2008). Glycogen synthase kinase 3 is a potential drug target for African trypanosomiasis therapy. *Antimicrobial Agents and Chemotherapy* 52, 3710–3717.

- Osmani, S. a, Pu, R.T., and Morris, N.R. (1988). Mitotic induction and maintenance by overexpression of a G2-specific gene that encodes a potential protein kinase. *Cell* 53, 237–244.
- Osolodkin, D.I., Zakharevich, N. V, Palyulin, V.A., Danilenko, V.N., and Zefirov, N.S. (2011). Bioinformatic analysis of glycogen synthase kinase 3: human versus parasite kinases. *Parasitology* 1–11.
- Ouna, B.A., Stewart, M., Helbig, C., and Clayton, C. (2012). The Trypanosoma brucei CCCH zinc finger proteins ZC3H12 and ZC3H13. *Molecular and Biochemical Parasitology* 1–5.
- O'regan, L., Blot, J., and Fry, A.M. (2007). Mitotic regulation by NIMA-related kinases. *Cell Division* 2, 25.
- Palenchar, J.B., and Bellofatto, V. (2006). Gene transcription in trypanosomes. *Molecular and Biochemical Parasitology* 146, 135–141.
- Parker, J.D.K., Bradley, B.A., Mooers, A.O., and Quarmby, L.M. (2007). Phylogenetic analysis of the Neks reveals early diversification of ciliary-cell cycle kinases. *PLoS One* 2, e1076.
- Parkinson, J.S. (2010). Signaling mechanisms of HAMP domains in chemoreceptors and sensor kinases. *Annual Review of Microbiology* 64, 101–122.
- Parsons, M., Valentine, M., and Carter, V. (1993). Protein kinases in divergent eukaryotes: identification of protein kinase activities regulated during trypanosome development. *Proceedings of the National Academy of Sciences of the United States of America* 90, 2656–2660.
- Parsons, M., Valentine, M., Deans, J., Schieven, G.L., and Ledbetter, J. a (1991). Distinct patterns of tyrosine phosphorylation during the life cycle of Trypanosoma brucei. *Molecular and Biochemical Parasitology* 45, 241–248.
- Parsons, M., Worthey, E.A., Ward, P.N., and Mottram, J.C. (2005a). Comparative analysis of the kinomes of three pathogenic trypanosomatids: Leishmania major, Trypanosoma brucei and Trypanosoma cruzi. 6, (15 September 2005).
- Parsons, M., Worthey, E.A., Ward, P.N., and Mottram, J.C. (2005b). Comparative analysis of the kinomes of three pathogenic trypanosomatids: Leishmania major, Trypanosoma brucei and Trypanosoma cruzi. 6, (15 September 2005).
- Peacock, L., Ferris, V., Bailey, M., and Gibson, W. (2009). Intraclonal mating occurs during tsetse transmission of Trypanosoma brucei. *Parasites & Vectors* 2, 43.
- Peacock, L., Ferris, V., Sharma, R., Sunter, J., Bailey, M., Carrington, M., and Gibson, W. (2011). Identification of the meiotic life cycle stage of Trypanosoma brucei in the tsetse fly. *Proceedings of the National Academy of Sciences of the United States of America* 1–6.
- Pearce, L.R., Komander, D., and Alessi, D.R. (2010). The nuts and bolts of AGC protein kinases. *Nature Reviews. Molecular Cell Biology* 11, 9–22.

- Peifer, C., and Alessi, D.R. (2008). Small-molecule inhibitors of PDK1. *ChemMedChem* 3, 1810–1838.
- Pradel, L.C., Bonhivers, M., Landrein, N., and Robinson, D.R. (2006). NIMA-related kinase TbNRKC is involved in basal body separation in *Trypanosoma brucei*. *Journal of Cell Science* 119, 1852–1863.
- Preußner, C., Jaé, N., and Bindereif, A. (2012). mRNA splicing in trypanosomes. *International Journal of Medical Microbiology* : IJMM 302, 221–224.
- Proto, W.R., Castanys-Munoz, E., Black, A., Tetley, L., Moss, C.X., Juliano, L., Coombs, G.H., and Mottram, J.C. (2011). *Trypanosoma brucei* metacaspase 4 is a pseudopeptidase and a virulence factor. *The Journal of Biological Chemistry* 286, 39914–39925.
- Proudfoot, C., and McCulloch, R. (2005). Distinct roles for two RAD51-related genes in *Trypanosoma brucei* antigenic variation. *Nucleic Acids Research* 33, 6906–6919.
- Pu, R.T., Xu, G., Wu, L., Vierula, J., O'Donnell, K., Ye, X.S., and Osmani, S.A. (1995). Isolation of a functional homolog of the cell cycle-specific NIMA protein kinase of *Aspergillus nidulans* and functional analysis of conserved residues. *The Journal of Biological Chemistry* 270, 18110–18116.
- Qi, M., and Elion, E. a (2005). MAP kinase pathways. *Journal of Cell Science* 118, 3569–3572.
- Quarmby, L.M., and Mahjoub, M.R. (2005). Caught Nek-ing: cilia and centrioles. *Journal of Cell Science* 118, 5161–5169.
- Queiroz, R., Benz, C., Fellenberg, K., Hoheisel, J.D., and Clayton, C. (2009). Transcriptome analysis of differentiating trypanosomes reveals the existence of multiple post-transcriptional regulons. *BMC Genomics* 10, 495.
- Raimondi, C., and Falasca, M. (2011). Targeting PDK1 in cancer. *Current Medicinal Chemistry* 18, 2763–2769.
- Redmond, S., Vadivelu, J., and Field, M.C. (2003). RNAit: an automated web-based tool for the selection of RNAi targets in *Trypanosoma brucei*. *Molecular and Biochemical Parasitology* 128, 115–118.
- Reininger, L., Tewari, R., Fennell, C., Holland, Z., Goldring, D., Ranford-Cartwright, L., Billker, O., and Doerig, C. (2009). An essential role for the *Plasmodium* Nek-2 Nima-related protein kinase in the sexual development of malaria parasites. *The Journal of Biological Chemistry* 284, 20858–20868.
- Reuner, B., Vassella, E., Yutzy, B., and Boshart, M. (1997). Cell density triggers slender to stumpy differentiation of *Trypanosoma brucei* bloodstream forms in culture. *Molecular and Biochemical Parasitology* 90, 269–280.
- Roditi, I., Schwarz, H., Pearson, T.W., Beecroft, R.P., Liu, M.K., Richardson, J.P., Bühring, H.J., Pleiss, J., Bülow, R., and Williams, R.O. (1989). Procyclin gene

expression and loss of the variant surface glycoprotein during differentiation of *Trypanosoma brucei*. *The Journal of Cell Biology* 108, 737–746.

Rokosz, L.L., Beasley, J.R., Carroll, C.D., Lin, T., Zhao, J., Appell, K.C., and Webb, M.L. (2008). Kinase inhibitors as drugs for chronic inflammatory and immunological diseases: progress and challenges. *Expert Opinion on Therapeutic Targets* 12, 883–903.

Roper, J.R., Guther, M.L.S., Milne, K.G., and Ferguson, M. a J. (2002). Galactose metabolism is essential for the African sleeping sickness parasite *Trypanosoma brucei*. *Proceedings of the National Academy of Sciences of the United States of America* 99, 5884–5889.

Rotureau, B., Subota, I., and Bastin, P. (2011). Molecular bases of cytoskeleton plasticity during the *Trypanosoma brucei* parasite cycle. *Cellular Microbiology* 13, 705–716.

Rotureau, B., Subota, I., Buisson, J., and Bastin, P. (2012). A new asymmetric division contributes to the continuous production of infective trypanosomes in the tsetse fly. *Development (Cambridge, England)* 1850, 1842–1850.

Roy, F., Laberge, G., Douziech, M., Ferland-McCollough, D., and Therrien, M. (2002). KSR is a scaffold required for activation of the ERK/MAPK module. *Genes & Development* 16, 427–438.

Rudenko, G. (2010). Epigenetics and transcriptional control in African trypanosomes. *Essays in Biochemistry* 48, 201–219.

Rudenko, G. (2011). African trypanosomes: the genome and adaptations for immune evasion. *Essays in Biochemistry* 51, 47–62.

Rusconi, F., Durand-Dubief, M., and Bastin, P. (2005). Functional complementation of RNA interference mutants in trypanosomes. *BMC Biotechnology* 5, 6.

Rätz, B., Iten, M., Grether-Bühler, Y., Kaminsky, R., and Brun, R. (1997). The Alamar Blue assay to determine drug sensitivity of African trypanosomes (*T.b. rhodesiense* and *T.b. gambiense*) in vitro. *Acta Tropica* 68, 139–147.

Scahill, M.D., Pastar, I., and Cross, G.A. (2008). CRE recombinase-based positive-negative selection systems for genetic manipulation in *Trypanosoma brucei*. *Molecular and Biochemical Parasitology* 157, 73–82.

Schnauffer, A., Domingo, G.J., and Stuart, K. (2002). Natural and induced dyskinetoplastic trypanosomatids: how to live without mitochondrial DNA. *International Journal for Parasitology* 32, 1071–1084.

Schneider, G. (2010). Virtual screening: an endless staircase? *Nature Reviews. Drug Discovery* 9, 273–276.

Schumann Burkard, G., Jutzi, P., and Roditi, I. (2011). Genome-wide RNAi screens in bloodstream form trypanosomes identify drug transporters. *Molecular and Biochemical Parasitology* 175, 91–94.

Schwede, A., and Carrington, M. (2010). Bloodstream form Trypanosome plasma membrane proteins: antigenic variation and invariant antigens. *Parasitology* 137, 2029–2039.

Scior, T., Bender, A., Tresadern, G., Medina-Franco, J.L., Martínez-Mayorga, K., Langer, T., Cuanalo-Contreras, K., and Agrafiotis, D.K. (2012). Recognizing Pitfalls in Virtual Screening: A Critical Review. *Journal of Chemical Information and Modeling*.

Seeliger, D., and De Groot, B.L. (2010). Ligand docking and binding site analysis with PyMOL and Autodock/Vina. *Journal of Computer-Aided Molecular Design* 24, 417–422.

Service, M. (2004). *Medical Entomology for Students* (Cambridge: Cambridge University Press).

Shi, H., Djikeng, A., Mark, T., Wirtz, E., Tschudi, C., and Ullu, E. (2000). Genetic interference in *Trypanosoma brucei* by heritable and inducible double-stranded RNA. *RNA (New York, N.Y.)* 6, 1069–1076.

Shi, H., Tschudi, C., and Ullu, E. (2006). An unusual Dicer-like1 protein fuels the RNA interference pathway in *Trypanosoma brucei*. *RNA (New York, N.Y.)* 12, 2063–2072.

Siegel, T.N., Hekstra, D.R., Wang, X., Dewell, S., and Cross, G. a M. (2010). Genome-wide analysis of mRNA abundance in two life-cycle stages of *Trypanosoma brucei* and identification of splicing and polyadenylation sites. *Nucleic Acids Research* 38, 4946–4957.

Simarro, P.P., Franco, J., Diarra, a, Postigo, J. a R., and Jannin, J. (2012). Update on field use of the available drugs for the chemotherapy of human African trypanosomiasis. *Parasitology* 139, 842–846.

Simpson, L., Aphasizhev, R., Gao, G., and Kang, X. (2004). Mitochondrial proteins and complexes in *Leishmania* and *Trypanosoma* involved in U-insertion/deletion RNA editing. *RNA (New York, N.Y.)* 10, 159–170.

Smith, A.J., Lauwaet, T., Davids, B.J., and Gillin, F.D. (2012). *Giardia lamblia* Nek1 and Nek2 kinases affect mitosis and excystation. *International Journal for Parasitology* 42, 411–419.

Steverding, D. (2008). The history of African trypanosomiasis. *Parasites & Vectors* 1, 3.

Steverding, D. (2010). The development of drugs for treatment of sleeping sickness: a historical review. *Parasites & Vectors* 3, 15.

Subramaniam, C., Veazey, P., Redmond, S., Hayes-Sinclair, J., Chambers, E., Carrington, M., Gull, K., Matthews, K., Horn, D., and Field, M.C. (2006). Chromosome-wide analysis of gene function by RNA interference in the african trypanosome. *Eukaryotic Cell* 5, 1539–1549.

- Sun, L., and Wang, C.C. (2011). The Structural Basis of Localizing Polo-Like Kinase to the Flagellum Attachment Zone in *Trypanosoma brucei*. *PloS One* 6, e27303.
- Szőor, B., Ruberto, I., Burchmore, R., and Matthews, K.R. (2010). A novel phosphatase cascade regulates differentiation in *Trypanosoma brucei* via a glycosomal signaling pathway. *Genes & Development* 24, 1306–1316.
- Szőor, B., Wilson, J., McElhinney, H., Taberner, L., and Matthews, K.R. (2006). Protein tyrosine phosphatase TbPTP1: A molecular switch controlling life cycle differentiation in trypanosomes. *The Journal of Cell Biology* 175, 293–303.
- Tanega, C., Shen, M., Mott, B.T., Thomas, C.J., MacArthur, R., Inglese, J., and Auld, D.S. (2009). Comparison of bioluminescent kinase assays using substrate depletion and product formation. *Assay and Drug Development Technologies* 7, 606–614.
- Tasker, M., Wilson, J., Sarkar, M., Hendriks, E., and Matthews, K. (2000). A novel selection regime for differentiation defects demonstrates an essential role for the stumpy form in the life cycle of the African trypanosome. *Molecular Biology of the Cell* 11, 1905–1917.
- Taylor, P., Blackburn, E., Sheng, Y.G., Harding, S., Hsin, K.-Y., Kan, D., Shave, S., and Walkinshaw, M.D. (2008). Ligand discovery and virtual screening using the program LIDAEUS. *British Journal of Pharmacology* 153 Suppl, S55–67.
- Tewari, R., Straschil, U., Bateman, A., Böhme, U., Cherevach, I., Gong, P., Pain, A., and Billker, O. (2010). The systematic functional analysis of Plasmodium protein kinases identifies essential regulators of mosquito transmission. *Cell Host & Microbe* 8, 377–387.
- Tibbetts, M.D., Shiozaki, E.N., Gu, L., McDonald, E.R., El-Deiry, W.S., and Shi, Y. (2004). Crystal structure of a FYVE-type zinc finger domain from the caspase regulator CARP2. *Structure (London, England : 1993)* 12, 2257–2263.
- Torkamani, A., and Schork, N.J. (2008). Prediction of cancer driver mutations in protein kinases. *Cancer Research* 68, 1675–1682.
- Torreele, E., Bourdin Trunz, B., Tweats, D., Kaiser, M., Brun, R., Mazué, G., Bray, M. a., and Pécol, B. (2010). Fexinidazole – A New Oral Nitroimidazole Drug Candidate Entering Clinical Development for the Treatment of Sleeping Sickness. *PLoS Neglected Tropical Diseases* 4, e923.
- Tu, X., and Wang, C.C. (2004). The involvement of two cdc2-related kinases (CRKs) in *Trypanosoma brucei* cell cycle regulation and the distinctive stage-specific phenotypes caused by CRK3 depletion. *The Journal of Biological Chemistry* 279, 20519–20528.
- Tu, X., and Wang, C.C. (2005). Pairwise knockdowns of cdc2-related kinases (CRKs) in *Trypanosoma brucei* identified the CRKs for G1/S and G2/M transitions and demonstrated distinctive cytokinetic regulations between two developmental stages of the organism. *Eukaryotic Cell* 4, 755–764.

- Ullu, E., Matthews, K.R., and Tschudi, C. (1993). Temporal order of RNA-processing reactions in trypanosomes: rapid trans splicing precedes polyadenylation of newly synthesized tubulin transcripts. *Molecular and Cellular Biology* 13, 720–725.
- Umeyama, T., and Wang, C.C. (2008). Polo-like kinase is expressed in S/G2/M phase and associated with the flagellum attachment zone in both procyclic and bloodstream forms of *Trypanosoma brucei*. *Eukaryotic Cell* 7, 1582–1590.
- Urbaniak, M.D. (2009). Casein kinase 1 isoform 2 is essential for bloodstream form *Trypanosoma brucei*. *Molecular and Biochemical Parasitology* 166, 183–185.
- Urbaniak, M.D., Guther, M.L.S., and Ferguson, M. a J. (2012a). Comparative SILAC Proteomic Analysis of *Trypanosoma brucei* Bloodstream and Procyclic Lifecycle Stages. *PLoS One* 7, e36619.
- Urbaniak, M.D., Mathieson, T., Bantscheff, M., Eberhard, D., Grimaldi, R., Miranda-Saavedra, D., Wyatt, P., Ferguson, M. a J., Frearson, J., and Drewes, G. (2012b). Chemical Proteomic Analysis Reveals the Drugability of the Kinome of *Trypanosoma brucei*. *ACS Chemical Biology*.
- Vanhollebeke, B., Uzureau, P., Monteyne, D., Pérez-Morga, D., and Pays, E. (2010). Cellular and molecular remodeling of the endocytic pathway during differentiation of *Trypanosoma brucei* bloodstream forms. *Eukaryotic Cell* 9, 1272–1282.
- Vassella, E., Krämer, R., Turner, C.M., Wankell, M., Modes, C., Van den Bogaard, M., and Boshart, M. (2001). Deletion of a novel protein kinase with PX and FYVE-related domains increases the rate of differentiation of *Trypanosoma brucei*. *Molecular Microbiology* 41, 33–46.
- Vassella, E., Probst, M., Schneider, A., Studer, E., Renggli, C.K., and Roditi, I. (2004). Expression of a major surface protein of *Trypanosoma brucei* insect forms is controlled by the activity of mitochondrial enzymes. *Molecular Biology of the Cell* 15, 3986–3993.
- Vassella, E., Reuner, B., Yutzy, B., and Boshart, M. (1997). Differentiation of African trypanosomes is controlled by a density sensing mechanism which signals cell cycle arrest via the cAMP pathway. *Journal of Cell Science* 110 (Pt 2, 2661–2671.
- Vickerman, K. (1965). Polymorphism and mitochondrial activity in sleeping sickness trypanosomes. *Nature* 208, 762–766.
- Vickerman, K. (1985). Developmental cycles and biology of pathogenic trypanosomes. *British Medical Bulletin* 41, 105–114.
- Walker, G., Dorrell, R.G., Schlacht, A., and Dacks, J.B. (2011a). Eukaryotic systematics: a user's guide for cell biologists and parasitologists. *Parasitology* 138, 1638–1663.
- Walker, R.G., Thomson, G., Malone, K., Nowicki, M.W., Brown, E., Blake, D.G., Turner, N.J., Walkinshaw, M.D., Grant, K.M., and Mottram, J.C. (2011b). High

throughput screens yield small molecule inhibitors of Leishmania CRK3:CYC6 cyclin-dependent kinase. *PLoS Neglected Tropical Diseases* 5, e1033.

Walker, R.G., Thomson, G., Malone, K., Nowicki, M.W., Brown, E., Blake, D.G., Turner, N.J., Walkinshaw, M.D., Grant, K.M., and Mottram, J.C. (2011c). High throughput screens yield small molecule inhibitors of Leishmania CRK3:CYC6 cyclin-dependent kinase. *PLoS Neglected Tropical Diseases* 5, e1033.

Wang, C.C. (1997). Validating targets for antiparasite chemotherapy. *Parasitology* 114 Suppl, S31–44.

Wang, X., Sheff, M. a, Simpson, D.M., and Elion, E. a (2011). Ste11p MEKK signals through HOG, mating, calcineurin and PKC pathways to regulate the FKS2 gene. *BMC Molecular Biology* 12, 51.

Wang, Z., Morris, J.C., Drew, M.E., and Englund, P.T. (2000). Inhibition of *Trypanosoma brucei* gene expression by RNA interference using an integratable vector with opposing T7 promoters. *The Journal of Biological Chemistry* 275, 40174–40179.

Wheeler, R.J., Gull, K., and Gluenz, E. (2012). Detailed interrogation of trypanosome cell biology via differential organelle staining and automated image analysis. *BMC Biology* 10, 1.

Wirtz, E., Leal, S., Ochatt, C., and Cross, G.A. (1999). A tightly regulated inducible expression system for conditional gene knock-outs and dominant-negative genetics in *Trypanosoma brucei*. *Molecular and Biochemical Parasitology* 99, 89–101.

Wloga, D., Camba, A., Rogowski, K., Manning, G., Jerka-Dziadosz, M., and Gaertig, J. (2006). Members of the NIMA-related kinase family promote disassembly of cilia by multiple mechanisms. *Molecular Biology of the Cell* 17, 2799–2810.

Wurst, M., Robles, A., Po, J., Luu, V.-D., Brems, S., Marentije, M., Stoitsova, S., Quijada, L., Hoheisel, J., Stewart, M., et al. (2009). An RNAi screen of the RRM-domain proteins of *Trypanosoma brucei*. *Molecular and Biochemical Parasitology* 163, 61–65.

Wyatt, P.G., Gilbert, I.H., Read, K.D., and Fairlamb, A.H. (2011). Target validation: linking target and chemical properties to desired product profile. *Current Topics in Medicinal Chemistry* 11, 1275–1283.

Zeytuni, N., and Zarivach, R. (2012). Structural and functional discussion of the tetra-trico-peptide repeat, a protein interaction module. *Structure (London, England : 1993)* 20, 397–405.

Zhang, J., Yang, P.L., and Gray, N.S. (2009). Targeting cancer with small molecule kinase inhibitors. *Nature Reviews. Cancer* 9, 28–39.

Zhao, Z., Lindsay, M.E., Roy Chowdhury, A., Robinson, D.R., and Englund, P.T. (2008). p166, a link between the trypanosome mitochondrial DNA and flagellum, mediates genome segregation. *The EMBO Journal* 27, 143–154.

Ziegelbauer, K., Quinten, M., Schwarz, H., Pearson, T.W., and Overath, P. (1990). Synchronous differentiation of *Trypanosoma brucei* from bloodstream to procyclic forms in vitro. *European Journal of Biochemistry / FEBS* 192, 373–378.